Modeling Approaches for Describing Microbial Population Heterogeneity

Lencastre Fernandes, Rita

Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

DTU Chemical Engineering
Department of Chemical and Biochemical Engineering

Modeling Approaches for Describing Microbial Population Heterogeneity

Rita Lencastre Fernandes

PhD Thesis

Kongens Lyngby, 2012

Technical University of Denmark
Department of Chemical and Biochemical Engineering
Center for Process Engineering and Technology
Søltofts Plads b. 229
DK-2800 Kongens Lyngby, Denmark
www.kt.dtu.dk

Abstract

Although microbial populations are typically described by averaged properties, individual cells present a certain degree of variability. Indeed, initially clonal microbial populations develop into heterogeneous populations, even when growing in a homogeneous environment.

A heterogeneous microbial population consists of cells in different states, and it implies a heterogeneous distribution of activities (e.g. respiration, product yield), including different responses to extracellular stimuli. The existence of a heterogeneous cell population may explain the lower productivities obtained for cultivations in large-scale reactors, where substrate and oxygen gradients are observed, in comparison to cultivations in well-mixed bench scale reactors.

Population balance models (PBM) have been used in a broad range of applications (e.g. crystallization, granulation, flocculation, polymerization processes) to predict distributions of certain population properties including particle size, mass or volume, and molecular weight. Similarly, PBM allow for a mathematical description of distributed cell properties within microbial populations. Cell total protein content distributions (a measure of cell mass) have been observed to provide a dynamic picture of the interplay between the cells and their surrounding extracellular environment.

The work here presented aimed at developing a model framework based on PBM as a tool to further understand the development of heterogeneous microbial populations subjected to varying environmental conditions. Three cases are presented and discussed in this thesis. Common to all is the use of *S. cerevisiae* as model organism, and the use of cell size and cell cycle position as single-cell descriptors.

The first case focuses on the experimental and mathematical description of a yeast population dynamics, in response to the substrate consumption observed

during batch cultivation. Cell size and cell cycle position distributions were used to describe the cell population. A two-stage PBM was developed and coupled to an unstructured model describing the extracellular environment. The good agreement between the proposed multi-scale model and experimental data (both the overall physiology and cell size and cell cycle distributions) indicates that a mechanistic model framework is a suitable tool for describing the microbial population dynamics in a bioreactor.

The second case provides an extension of the proposed model framework (PBM coupled to an unstructured model) to a continuous cultivation. A compartment model approach was applied for addressing situations where two zones (compartments) are formed due to non-ideal mixing in the bioreactor. In particular, this approach was used in order to assess the impact of the degree of compartmentalization (i.e. deviation from the ideal mixing case) on the population dynamics and overall system performance under various operation conditions (substrate feed concentration and dilution rate). It was possible to conclude that the deviation from ideal mixing may have a significant effect on the observed system dynamics. Moreover, oscillatory pseudo-steady states may be observed for particular combinations of operating conditions and degree of compartmentalization.

In the third study attention was paid to the integration of the proposed model framework in a computational (CFD) fluid dynamic model. The anaerobic growth of a budding yeast population in a continuously run microbioreactor was used as example. The proposed integrated model describes the fluid flow, the local cell size and cell cycle position distributions, as well as the local concentrations of glucose, ethanol and biomass throughout the reactor. This work has proven that the integration of CFD and population balance models, for describing the growth of a microbial population in a spatially heterogeneous reactor, is feasible, and that valuable insight on the interplay between flow and the dynamics of a budding yeast population (e.g. formation of substrate gradients and non-growth zones) is gained. *In silico* simulation tools, as the one proposed, may be used for hypothesis generation and testing, and when coupled to an experimental set-up may be used for process and reactor design optimization.

Resumé

Mikrobielle populationer typisk er beskrevet af gennemsnitlige egenskaber, men individuelle celler viser en vis grad af variation. Mikrobielle populationer, der starter som klonale udvikler sig til heterogene populationer, selv når de gror i et homogent miljø.

En heterogen mikrobiel population består af celler i forskellige tilstande og det indebærer en heterogen distribution af aktiviteter (bl.a. respiration, produkt udbytte), inklusiv forskellige responser af ekstracellulære stimuli. Eksistensen af heterogene celle populationer kan forklare de lave produktiviteter konstateret i stor skala kulturer, hvor substrat- og iltgradienter er observeret, sammenlignet med kulturer i velblandede laboratorieskala reaktorer. Population balance modeller (PBM) bruges i en bred række af anvendelser (f.eks. krystallisering, granulering, flokkulering og polymerisation) til at forudse distributioner af visse egenskaber, blandt andet partikelstørrelse, masse, volumen og molekylær vægt. Ligeledes tillader PBM en matematisk beskrivelse af distribuerede celle egenskaber af mikrobielle populationer. Distributioner af cellers totale proteinindhold (en måling af celle masse) bruges for at danne et dynamisk billede af samspillet mellem cellerne og deres omgivende ekstracellulært miljø.

Dette PhD projekt stiler efter at udvikle et model skelet baseret på PBM som et værktøj til nærmere at forstå udviklingen af heterogene mikrobielle populationer under variable miljøtilstande. Tre cases præsenteres og diskuteres i denne afhandling, hvor *S. cerevisiae* bruges som modelorganisme, og hvor de enkelte celler beskrives ud fra cellestørrelsen og cellecyklus positionen.

I den første case fokuseres på den eksperimentelle og matematiske beskrivelse af gær populations dynamik, i forhold til substratforbruget i løbet af en batch kultur. Cellestørrelsens og cellecyklus positionens distributioner bruges til at beskrive celle populationen. En to-stadie PBM udvikles og kobles sammen med

en ustruktureret model som beskriver det ekstracellulære miljø. Den gode overensstemmelse mellem den foreslåde multi-skala model og de eksperimentelle data (både den overordnede fysiologi samt distributionerne af cellestørrelsen og cellecyklus positionen) er tegn på, at den mekanistiske model er et passende værktøj til at beskrive dynamikken af mikrobielle populationer i en bioreaktor.

Den anden case leverer en videreudvikling af det foreslået model skelet (PBM koblet sammen med en ustruktureret model) til en kontinuert kultur. I casen benyttes *compartment model approach* i det tilfælde hvor to zoner (rum) dannes pga. ikke-ideel opblanding i bioreaktoren. Denne fremgangsmåde bruges til at evaluere hvordan graden af rumopdeling påvirker populationsdynamikkerne og den overordnede systempræstation, under forskellige operations vilkår (substrat foderkoncentration og fortyndingshastighed). Det var muligt at konkludere at afvigelser fra ideel opblanding kan have en betydelig effekt på systemdynamikkerne. Derudover kunne det ligeledes konkluderes at særlige kombinationer af operations vilkår og graden af rumopdeling resulterer i oscillatoriske pseudostationære tilstande.

Den tredje case fokuserer på integrationen af det foreslået model skelet i en *Computational Fluid Dynamics* (CFD) model. Anaerobisk vækst af en gær population i en kontinuert mikrobioreaktor bruges som eksempel. Den foreslået integreret model beskriver væskestrømsfeltet, lokale distributioner af cellestørrelse og cellecyklus position samt lokale koncentrationer af glukose, ethanol, og biomasse over hele reaktoren. Denne case viste at integrationen af PBM i CFD er i stand til at beskrive væksten af mikrobielle populationer, og giver et værdifuldt indblik i samspillet mellem væskestrøm og dynamikkerne af gær populationer (f. eks. dannelse af substrat gradienter og vækstfri zoner). Udviklingen af *in silico* simulationsværktøjer kan bruges til at generere og teste hypoteser, og koblet sammen med en eksperimentel opstilling kan disse bruges til optimering af proces og reaktor design.

Preface

This thesis was prepared at the Department of Chemical and Biochemical Engineering (KT), at the Technical University of Denmark (DTU) in partial fulfillment of the requirements for acquiring the Ph.D. degree in engineering.

The work here presented was developed in the period from November 2009 until early December 2012, and it was partly funded by DTU and partly funded by the Danish Council for Strategic Research, as part of a larger research project entitled *Towards robust fermentation processes by targeting population heterogeneity at microscale* (project no. 09-065160). Furthermore, collaborations with the BIOMATH group at Ghent University and the Center for Microbial Biotechnology (CMB) at DTU System Biology were established under the frame of the ERA-IB project *Targeting population heterogeneity at microscale for robust fermentation processes* (project no. EIB.08.031).

Associated Professor Krist V. Gernaey (DTU Chemical Engineering) was the principal supervisor for the project, with Professor Anker D. Jensen (DTU Chemical Engineering) and Associated Professor Ingmar Nopens (Ghent University) as co-supervisors. The data analysis and modeling research work was performed at the Center for Process Engineering and Technology (Process, DTU Chemical Engineering), while the experimental work was performed at CMB (DTU System Biology). The preparation of a review article and familiarization with the formulation and solution methods for population balance models was initiated during a two month external stay at Ghent University in 2010.

At the beginning of the project a main idea was on the table: it had been suggested that non-ideal mixing was the probable cause for deviating behaviors of large-scale cultivations relatively to bench-scale, as the existence of different microenvironments found within a large-scale bioreactors resulted in broader distributions of cell behaviors (rather than identical), i.e. in the existence of heterogeneous cell populations. Starting from the question - how could population balance

models, which are able to describe distributed properties, and computational fluid dynamics, be used to simulate the dynamics of heterogeneous microbial populations in response to variation in the extracellular environment? - the project slowly developed into the work summarized in this thesis.

In this thesis different modeling approaches are discussed in order to address the challenge of describing the dynamics of a microbial population as a result of the interplay of cells and the extracellular environment. Although a microbial population may be described by a multiplicity of distributed cell phenotypes, in this thesis, focus was set on describing cell size and cell cycle position distributions (due to link between these atributes and cell growth). The coupling between the population and extracellular environment (in terms of substrate availability and consumption) was particularly addressed, in order to attain a model framework that could be used for exploring the interplay between the microbial cells and their surrounding environment.

During the time of the project, in a collaboration with CMB (DTU System Biology), additional efforts were made in order to describe the dynamics of green fluorescent protein distributions for a reporter strain cultivation. In this case, a *S. cerevisiae* strain where the expression of green fluorescent protein (GFP) is controlled by a promoter for a ribosomal protein had been constructed, and experimental studies were being conducted at CMB. Although the experimental efforts and data analysis yielded interesting qualitative results (see [1]), the gained insight was not sufficient for establishing a model that captured, even if only qualitatively, the experimentally observed behaviors. Therefore, this *subproject* was not included in this thesis.

Parallel, at Ghent University, a computational fluid dynamics model of a stirred tank reactor was implemented with the aim of studying the non-ideal mixing at different scales, in particular the occurence of zoning and compartmentalization within the bioreactor. The objective was to incorporate a kinetic model based on average variables (unstructed model) in a first stage, and a distributed model (PBM) at a later stage, so the impact of non-ideal mixing on the biological system was assessed. In collaboration with Ghent University, experimental mixing tests for a coarse validation of the CFD model were performed, at DTU, as part of this PhD project. Unfortunately, the integration of CFD and kinectic (unstructured or distributed) models has not been achieved at this point, and it is, thus,

not included in this thesis.

Despite these less successful points (which are part of nearly all research projects, but are seldom reported), at the end of this PhD project, I believe that the work summarized in this thesis consists of small (but relevant) step towards an understanding of the interplay between cell populations and the extracellular environment, not only from a qualitative but also quantitative perspective. Indeed, I believe the modeling approaches presented in this thesis not only prove that population balance models can be used for describing realistically the dynamics of a microbial population, but also demonstrate that the PBM framework can be useful for identifying the knowledge gaps and for generating hypotheses and designing elucidative experiments.

Kgs. Lyngby, December 7th, 2012

Rita Lencastre Fernandes

Acknowledgements

I would like to thank my three supervisors - Krist V. Gernaey, Ingmar Nopens and Anker D. Jensen - as well as to Anna Eliasson Lantz (Assoc. Professor at CMB), for their valuable help and for the trust in my efforts, specially in the moments I was not sure I would succeed. I am, in particular, grateful to Krist for believing in me and accepting to supervise by MSc. thesis project, later offering me the opportunity to return to DTU for pre-PhD project, that then became the PhD project I am now concluding.

I am also specially thankful to Magnus Carlquist (a sort of big brother in the population heterogeneity projects) for all the help and energy in the fermentation work, as well as for all the discussions about data analysis and interpretation. I greatly appreciated his availability and assistance, even after moving to a challenging research position in Lund. Furthermore, I would like to thank Anna-Lena Heins (CMB) for helping out with cultivations and for being a confident in all the frustrating moments any PhD project includes.

I thank also all the people that work (and have worked) in Process for all the fun moments and friendship, and patience for hearing 101 times *tired*, as my reply to the daily *how are you?*. I leave a special thank you note to my office mate Rui for forgiving all the times I possibly have cursed the computer for being slow, and myself for making idiotic mistakes.

Finally, I thank my family for being a constant in my life, and my friends - Teresa, Joana, Inés, Chus, Signe, Rita, Watson, Trond, Xico - for making my days fun and being my family outside Portugal. Last but not least, I leave an enormous *thank you* to Simon for being beside me in the happy and sad moments of the last three years.

Contents

List of Figures

List of Tables

Abbreviations

ATP	Adenosine triphosphate
CFD	Computational Fluid Dynamics
DNA	Deoxyribonucleic acid
FCM	Flow cytometry
FDA	Food and Drug Administration
FITC	Fluoresceine-isothiocyanate
FSC	Forward scatter signal
G1	Cell cycle phase
G2	Cell cycle phase
M	Mitosis (cell cycle phase)
NAPH	Nicotinamide adenine dinucleotide
PBE	Population balance equation(s)
PBM	Population balance model(s)
PI	Propidium iodide
RNA	Ribonucleic acid
S	Cell cycle phase
Start	Cell cycle regulation point between G1 and S phases
SSC	Side scatter signal

Chapter 1

Introduction

1.1 Heterogeneity in microbial cultivations

Microbial fermentations or cultivations can be generally described as the growth of a microorganism strain in a controlled reactor. The main product may be the biomass itself (e.g. production of *Saccharomyces cerevisiae* for the baking industry) or a molecule resulting from the cell metabolism (e.g. production of enzymes or other proteins). In most industrial fermentation processes a single strain is used, although few examples of mixtures of more than one microorganism can be found (e.g. cultures of *Lactobacillus* and *Streptococcus* used for yoghurt production). In this thesis, only single strain cultivations will be addressed.

As the microbial cultures used in the bioprocess are, most commonly, isogenic, cell heterogeneity would not be expected if the environment were well controlled at adequate conditions. Nonetheless, population heterogeneity has been observed, particularly at large scale. In fact, the scale up of bioprocesses

from bench-scale to large scale may lead to lower yields and productivities and an increased by-product formation [2, 3].

Indeed, whilst the assumption of a perfectly mixed reactor might be realistic for bench-scale reactors, it certainly is not for large scale bioreactors. Due to limited mixing and mass transfer, gradients of, for example, substrate, oxygen and pH are observed in larger reactors [4]. In fact, substrate concentrations may range from high concentrations close to the feed port to residual concentrations in zones more distant to this port, the latter caused by different rates of mixing and biological reaction [5]. Cells circulating in the reactor are subjected to successively changing conditions, which, by inducing genetic, metabolic and physiological responses, are held responsible for the development of heterogeneous populations.

Due to the observed decrease in performance at large scale, heterogeneity in bioprocesses was felt to be undesirable. Nonetheless, it might be the key to cell population robustness as observed in tumors [6] or in cases of bacterial persistence [7]. Similar to mechanical stress, which can be exploited to control fungal morphologies to increase overall productivity [8], it might also be possible to take advantage of heterogeneity in a microbial population for process optimization.

In fact, to understand and harness cell heterogeneity may show us a new path for achieving improved robustness in bioprocesses. This, in fact, is also the central hypothesis in the project 'Towards robust fermentation processes by targeting population heterogeneity at microscale', which this PhD project is part of (see Preface). Indeed, it is believed that modeling can play a central role in identifying and describing the main phenomena influencing population heterogeneity.

The monitoring and control of bioprocesses, found in industry today, does not account for the heterogeneity in microbial populations. The cell properties, determined using on-line, at-line or off-line monitoring methods, correspond to averaged values and, thus, camouflage valuable information on the dynamics of the population (Figure 1.1).

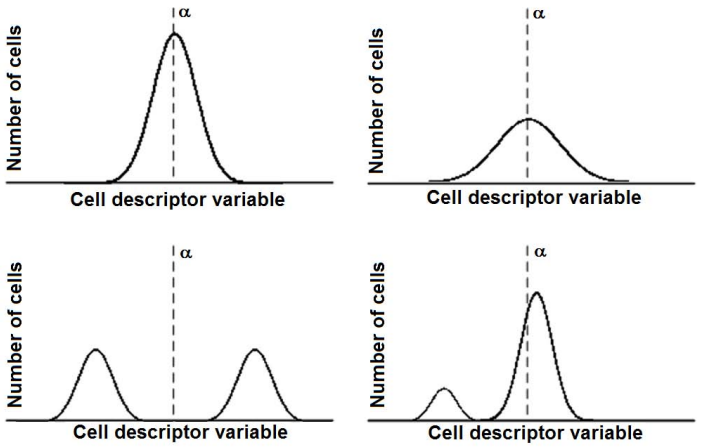

Figure 1.1: Schematic description of distributed cell properties. The structure of a population is masked by the use of average cell properties. Different population distributions can correspond to the same mean value α of an experimentally quantifiable cell parameter (e.g. DNA content, NAD(H) concentration, cell mass) (Based on [9]).

Different types of models have been proposed for fermentation processes [10]. Most often mechanistic unstructured models based on stoichiometry and Monod-type expressions are used. In this type of models, cell metabolites and pathways are not described in detail. With the advent of metabolic flux analysis, structured models describing various metabolic pathways and in some cases regulatory loops, in the form of kinetic equations and stoichiometric balances, have been proposed. Whether unstructured or structured, these models describe the behavior of an average cell, considering implicitly that all cells in a cultivation are alike and behave in a similar fashion.

In an effort to understand the occurrence and development of heterogeneous microbial populations, several experimental studies have been reported in recent years, as the number of experimental methods available for single-cell analysis has boomed [11]. However, this knowledge has not yet been integrated into a generally accepted modeling framework that is able to account for distributed properties within a cell population, and thus can be used in the design and control

of bioprocesses [12].

In the last decade, a great development of experimental methods for single cell analysis has taken place and various experimental studies have shown that phenotypic heterogeneity is ubiquitous. In the case where the cell population is subjected to variations in growth conditions, for example when circulating in a large fermentor where mixing is not ideal, a higher degree of cell heterogeneity is observed. The consequent existence of a broader distribution of cell behaviors has been pointed out as the a potential explanation to the difficult scalability that is sometimes observed when using kinetic information obtained from lab-scale experiments, as well as models validated using lab-scale data, to predict fermentation profiles and overall performance in larger scales [3].

Since the 1970's population balance models (PBM) have been proposed for describing microbial populations [13]. Population balance models are partial-differential equations, and their formulation and solution is not straightforward as the cases where unstructured or even structured models are used. This is reflected in the scarce number of studies discussing PBM formulation and predictions in the light of experimental evidences. In fact, even in areas in which the use of PBM is more established, such as crystallization and flocculation, a research trend towards linking experimental data to the theoretical modeling work has been observed recently [14].

PBM may thus form an interesting modeling tool to account for cell heterogeneity and serve as investigation tool to understand the dynamics of a cell population. Moreover, computational fluid dynamic models may be used to describe the spatial gradients of substrate, pH etc observed in non-ideal mixed reactors. By integrating the PBM with a CFD, the interplay between varying extracellular cellular conditions and the dynamic cell population may be further explored.

1.2 Research hypothesis

The work presented in this thesis relies on two key hypotheses:

- PBM provides an adequate mathematical framework for describing dynamic distributions of cell properties as functions of the varying extracellular environment.

- The dynamics of a microbial population under a varying extracellular environment observed for lab-scale cultivations is similar to the dynamics observed in larger scale, for the same environment variations. Therefore, establishing models based on behaviors observed in lab-scale can be used to gain an understanding of the development of heterogeneous cell populations, which can in the future be translated to larger scales.

1.3 Scope of the work and specific research goals

Generally, this PhD project aimed at developing a modeling framework able to predict the dynamics of heterogeneous microbial populations in response to varying extracellular conditions. The work presented focuses on population heterogeneity in single-strain cultivations of *Saccharomyces cerevisiae*. The following specific research goals were addressed:

- To formulate and solve a PBM model for a lab-scale batch cultivation using cell properties that are deeply linked to growth conditions (such as cell size and cell cycle position) as the PBM variables. The formulation of the model kernels as functions of the substrate availability is based a thorough analysis of experimental cell physiology data and single cell measurements

(obtained by using flow cytometry), as well as available literature on yeast *omics* and physiology.

- To extend the PBM model framework to continuous cultivations where spatial heterogeneities are observed in the bioreactor.

- To develop a simulation tool, as proof-of-concept of the integration of PBM and CFD for describing local distributions of cell properties, as well as local extracellular environments, in a non-ideally mixed bioreactor.

1.4 Thesis outline

This thesis consists of 7 main chapters:

- *Chapter 2* provides an introduction to *S. cerevisiae* as an important work horse organism in both research and industrial contexts. A first part consists of a brief description of budding yeast physiology and the cultivation modes traditionally used in research and industrial contexts. A second part provides a short review on mathematical models proposed in the scientific literature for describing aerobic budding yeast cultivations, based on average behaviors and measurements.

- *Chapter 3* consists of an overview on experimental methods for single cell analysis and provides more details on the use of flow cytometry for single-cell analysis of microbial samples, including a section on analysis of the collected flow cytometric data.

- *Chapter 4* consists of an overview of modeling techniques for description of microbial populations, i.e. that account for distributed cell properties. A first part provides a description of population balance models and a literature review on the use of PBM for describing microbial cultivations. A second

part provides an overview on the use of computational fluid dynamics (CFD) for describing non-ideally mixed reactors, and the potential of coupling CFD and PBM for simulating the behavior of cell populations in spatially heterogeneous bioreactors.

- *Chapter 5* presents the development of a PBM for describing the dynamics of cell size and cell cycle position of a budding yeast population, in response to the substrate consumption during batch cultivation. The PBM is further coupled to an unstructured model describing the substrate and metabolite concentrations in the cultivation media. Experimental physiology and single cell data are reported and used in the formulation of the models.

- *Chapter 6* focuses on the development of microbial populations for continuous cultivations in a spatially heterogeneous stirred tank bioreactor. A compartment model approach is used in order to assess the effect of reactor compartmentalization on the population structure and system behavior in comparison to the case of an ideally mixed bioreactor (a single compartment model).

- *Chapter 7* provides a proof-of-concept for the integration of PBM and CFD for describing population dynamics (cell size and cell cycle distributions) in spatially heterogeneous bioreactor. The anaerobic growth of a budding yeast population in a continuously run microbioreactor was modeled. Local substrate concentration and local cell size and cell cycle distributions were determined and compared in order to evaluate the effect of the feed flow rate on the population structure.

- *Chapter 8* concludes the thesis and provides a discussion on the future challenges and perspectives.

1.5 Publications included in the thesis

The following three accepted publications have resulted from the work presented in this thesis, and published manuscripts are provided in Appendix A:

- Lencastre Fernandes R, Nierychlo M, Lundin L, Pedersen AE, Puentes Tellez PE, Dutta A, Carlquist M, Bolic A, Schäpper D, Brunetti AC, Helmark S, Heins A-L, Jensen AD, Nopens I, Rottwitt K, Szita N, van Elsas JD, Nielsen PH, Martinussen J, Sørensen SJ, Lantz AE, Gernaey KV. *Experimental methods and modeling techniques for description of cell population heterogeneity*. Biotechnology Advances (2011) 29:575-599.

Parts of this review article are reproduced in Section 1.1 of this Chapter, as well as included in Chapter 3 and Chapter 4. References to more recent studies have been added where suitable.

- Lencastre Fernandes R, Carlquist M, Lundin L, Heins A.-L., Dutta A , Sørensen SJ, Jensen AD, Nopens I, Eliasson Lantz A, Gernaey KV. *Cell mass and cell cycle dynamics of an asynchronous budding yeast population: experimental observations, flow cytometry data analysis and multi-scale modeling*. Biotechnology and Bioengineering. (2012) (DOI: 10.1002/bit.24749)

Most of this paper is reproduced in Chapter 5. Small adaptations were made in order to avoid repetitions and improve the readibility of this thesis. A section providing details on the discretization and implementation of the population balance model was included.

- Lencastre Fernandes R, Krühne U, Nopens I, Jensen AD, Gernaey KV. *Multi-scale modeling for prediction of distributed cellular properties in response to substrate spatial gradients in a continuously run microreactor*, In: Iftekhar A. Karimi and Rajagopalan Srinivasan, Editor(s), Computer Aided Chemical Engineering, Elsevier (2012) 31:545-549

Chapter 7 provides an extended version of this short conference contribution. Details on model implementation and integration of the population balance model

into the computational fluid dynamics model were included, as well as further results and discussion.

1.6 Other publications

During the 3 years PhD project, and a prior 3 month pre-PhD project, other collaborations were pursued and resulted in the following publications, which have not been included in this thesis:

- Lencastre Fernandes R, Bodla V, Carlquist M, Heins AL, Eliasson Lantz A, Sin G, Gernaey KV. *Applying mechanistic models in bioprocess development*. In: Carl-Fredrik Mandenius and Nigel Titchener-Hooker, Editors, Advances in Biochemical Engineering/Biotechnology, Springer (2012) (accepted for publication)
- Carlquist M, Lencastre Fernandes R, Helmark K, Heins AL, Gernaey KV, Eliasson Lantz A. *Physiological heterogeneities in microbial populations and implications for physical stress tolerance*. Microbial Cell Factories (2012) 11:94
- Schäpper D, Fernandes RL, Lantz AE, Okkels F, Bruus H, Gernaey KV. *Topology Optimized Microbioreactors*. Biotechnology and Bioengineering. (2011) 108:786-796

Chapter 2

Saccharomyces cerevisiae as microorganism of interest

The work presented in this thesis focuses on describing the growth of microbial populations during controlled cultivations in bioreactors. *Saccharomyces cerevisiae* has been selected as the microorganism to be studied, due to the vast amount of prior knowledge of *S. cerevisiae* physiology, as well as experimental know-how, of collaboration partners involved in the larger research project, this PhD was part of (see Preface).

This chapter provides a brief overview of *S. cerevisiae* as model organism in research and as an industrial relevant strain. Furthermore, a short description of key aspects of submerged (or liquid) cultivations and monitoring tools is presented, followed by a succinct description of general features of yeast physiology and cell cycle. Finally, a short overview of modeling approaches often used for describing microbial fermentations, based on averaged cell behaviors, is provided.

2.1 *S. cerevisiae* as research model organism

S. cerevisiae is one of the most studied microorganisms and it is considered to be a model organism, i.e. it is regarded as representative of a larger class of living beings when studying a particular process or phenomenon (e.g. [15, 16, 17]). In the specific case of *S. cerevisiae*, its use as model system, in particular for molecular genetics research, relies on the fact that basic cellular mechanics of replication, recombination, cell division and metabolism are generally conserved between yeast and larger eukaryotes, including mammals. Consequently, it is not surprising that the complete genome of *S. cerevisiae* was already published in 1996 [18] - the sequence of the human genome was reported in 2001 [19]. This led to extensive work on the linking of genes to function, by proposing metabolic and functional maps. A comprehensive community resource (Saccharomyces Genome Database) gathering among others gene annotations, genetic and physical analyses and genome-wide analysis tools has been available since 1998 [20].

2.2 *S. cerevisiae* as an industrially relevant organism

Saccharomyces cerevisae has been used for thousands of years in traditional processes as bread baking, wine making and beer brewing, and it is thus also designated as baker's or brewer's yeast. From an industrial point of view many other industrial processes have been established using *S. cerevisiae* as microorganism, besides the traditional baking, brewing and winemaking.

The production of yeast for the baking industry or in the form yeast extract (e.g. used as food additive), as well as industrial fermentation processes for biofuel production (e.g. bioethanol from sugar and starch feedstocks) have been established for decades [21]. Other established industrial processes for production of glycerol, pyruvate, and organic acids (which can be used as building blocks in the production of active pharmaceutical ingredients) also use *S. cerevisiae* as

fermenting organism.

Moreover, *S. cerevisiae* is used for the production of several heterologous (pharmaceutical) proteins (e.g. insulin and vaccines). Indeed, most of the pharmaceutical proteins produced by microbial eukaryotic cells that have been approved by the EMEA or FDA are produced almost exclusively in *S. cerevisiae* [22]. In comparison to other microbial eukaryotes, *S. cerevisiae* presents some advantages with regard to its application in new processes: (i) a large knowledge base has been accumulated, and the authorities (e.g. FDA) approval process is often easier (and thus cheaper) than when using a new unknown organism; (ii) *S. cerevisiae* has the status of Generally Regarded as Safe (GRAS) organism.

Reviews of therapeutical proteins produced in yeast has been made by Gerngross [23] and Walsh [24]. Although mammalian cell lines, due to their the ability to perform post- translational modifications, are used for production of the largest share of new biopharmaceuticals (in particular the ones approved in the last 2-3 years) [25], it is not expected that *S. cerevisiae* will stop being used as work horse both in industrial and research environments. Indeed, future applications of metabolic engineered strains of *S. cerevisiae* for production of pharmaceuticals (e.g. [26]), biofuels (e.g. [27]), and other bulk chemicals [28] are being discussed and research work is frequently reported.

2.3 Fermentation process and cultivation modes

Generally, a microbial cultivation can be described as the growth of a microorganism in a controlled environment. Cultivations may be submerged, where cells are suspended in a liquid cultivation medium, or in solid-state where the solid culture substrate (e.g. wheat) may, for example, be deposited in flat trays and seeded with microorganisms. In this work, only submerged fermentations will be addressed. A review on recent advances in solid-state fermentation may be found elsewhere (e.g. [29, 30]).

In a submerged cultivation, temperature, pH, agitation and aeration (gas flow) are typically pre-defined operating conditions and kept at the set point values using control loops. The cultivation medium typically contains the carbon and nitrogen sources, as well as trace metals and ions (e.g. phosphate) necessary for the microorganisms growth. A schematic representation of a fermentation laboratory set-up is presented in Figure 2.1.

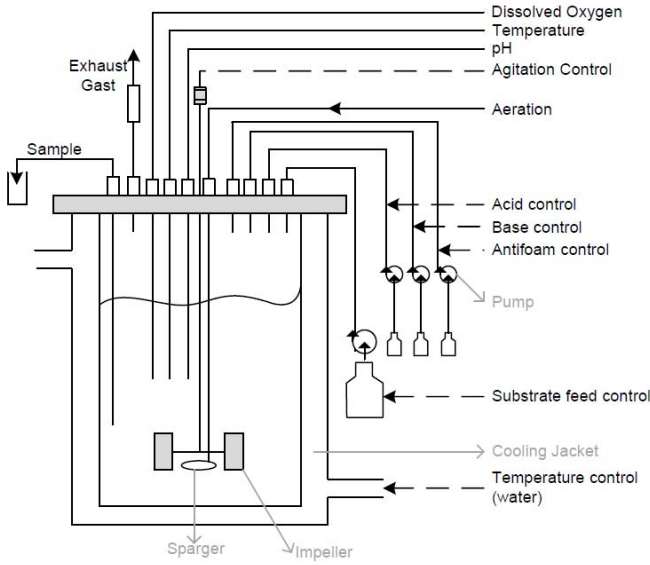

Figure 2.1: Schematic representation of a laboratory-scale bioreactor for aerobic fermentations (based on [31]).

A submerged fermentation may be run in three different cultivation modes. The choice of the method depends on the desired product and the organism being used, as well as other factors such as available equipment, desired product titer and amount, and other aspects influencing process economics. The three cultivation modes can be described as follows:

- **Batch**: in this mode, the reactor is a nearly closed system. All media compo-

nents are present from the beginning and there is no liquid inflow or outflow. The reactor is, however, constantly aerated and the exhaust gas flows out of the reactor. Cells experience a high glucose concentration in the beginning of the cultivation, and, in the case of *S. cerevisiae*, this results in the production of ethanol. Cells grow, thus, at their maximum specific growth rate until one of the nutrients becomes limiting. The initial high concentration of nutrients may cause osmotic stress to the cells.

- **Continuous**: after an initial batch phase for biomass production, the cultivation mode may be switched to continuous (also known as chemostat). In this mode, fresh medium is added at the reactor inlet and cultivation broth is collected at the outlet, at the same flow rate. This results in a constant volume, as well as constant concentrations (of e.g. substrate, metabolites) in the reactor once steady state is reached. One nutrient will be the limiting one and the rate of addition of this nutrient determines the specific growth rate of the cells.

- **Fed-batch**: this cultivation mode is started after an initial batch phase where biomass is produced. In the fed-batch cultivation, fresh medium is added to the reactor, but no liquid outflow is collected. The cultivation volume, thus, increases. Growth is limited by one nutrient, and the feed flow of this nutrient determines the specific growth rate. The feed rate profiles may be imposed (linear or exponential) or result from a control loop based on, for example, oxygen limitation or heat transfer limits.

All three cultivation modes are used at industrial scale, although fed-batch cultivations are often preferred, particularly in the case of protein production [32], as higher volumetric productivity can be achieved. In a research context, batch and continuous cultivations may be preferred, as the time and volume variation observed for fed-batch poses poses additional challenges regarding the interpretation of culture performance and metabolic responses. Knowledge extracted from continuous cultures at different growth rates (while other conditions are kept con-

stant) can, for example, be applied in the development and design of a corresponding fed-batch process.

2.4 Monitoring of *S. cerevisiae* cultivations

Monitoring of cultivations can be defined as the measurement of various variables in order to follow the process, and allow for control of the desired cultivation conditions. Monitoring methods are typically classified as off-line, at-line and on-line depending on the location where the measurement is performed, and time frame for obtaining the analysis results:

- **Off-line**: a sample is collected from the bioreactor and the analysis is performed in a different location (e.g. analytical laboratory). The time interval between sampling and release of the analysis results may be of hours up to several days.

- **At-line**: the collected sample is analyzed in the vicinity of the bioreactor and the analysis results are obtained in a time range of minutes to hours.

- **On-line**: the analysis relies on a fully automatic procedure. This type of measurements may be performed *in situ* (e.g. using a probe placed inside the bioreactor) or *ex situ* (e.g. a sample is automatically collected from the bioreactor and transported to a flow cell where the measurement is performed). The time frame for this type of analysis is of seconds to minutes. If the aim is to use the analysis for control of the process, time-delays should be avoided as real-time measurement and control is desirable.

Several physical and chemical variables are typically measured on-line in industrial and research cultivations including feed flow rate, temperature, pH, aera-

tion gas flow (rate and composition), electrical power consumption, stirrer speed, pressure, dissolved oxygen concentration (for aerobic cultivations), pH and exhaust gas composition [33]. For example, the control of temperature is usually based on the water flow in the cooling jacket, whereas pH control is achieved by addition of base (e.g. $NaOH$ or KOH solutions) or acid (e.g. H_2SO_4 orHCl).

2.5 Brief notes on *S. cerevisiae* physiology and cell cycle

S. cerevisiae is a Crabtree positive organism: in the presence of high gluocse concentrations, its respiratory capacity is exceeded and the glucose is additionally fermented to ethanol (as schematically illustrated in Figure 2.2). When glucose is depleted, ethanol is used as carbon source and consumed by oxidation. A typical

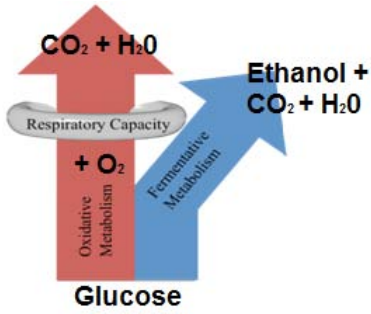

Figure 2.2: Schematic representtion of the Crabtree effect observed for *S. cerevisiae*.

batch cultivation consists of four phases. After an (1) initial lag-phase, a (2) first growth phase on glucose (and corresponding ethanol accumulation) takes place. Upon glucose depletion, cells undergo a metabolic rearragement during a short period called (3) diauxic shift. A (4) second growth phase, where the accumulated ethanol is used as carbon source, follows. When ethanol is depleted cells enter a stationary phase. Also other metabolites like glycerol and acetate are accumulated during the growth phase on glucose, and later used for growth (in the later

ethanol growth phase), though in significantly lower concentrations than ethanol.

In the cases of continuous and fed-batch cultivation, the fresh medium feed is often started when glucose is near depletion, or in the fed-batch case when, for example, the dissolved oxygen concentration has reached a certain set-point value and this variable will be used to control the substrate feed.

The yeast reproduction is based on the asexual process of budding (hence, *S. cerevisiae* is also known as budding yeast) in which a new daughter cell forms of an outgrowth of the original cell. The cell cycle normally consists of the phase G1, S, G2 and M (Figure 2.3). When the regulation START point is passed, cells enter the S-phase and the bud formation as well as the DNA replication are initiated. When cells enter the G2 pahse, DNA replication has been completed.

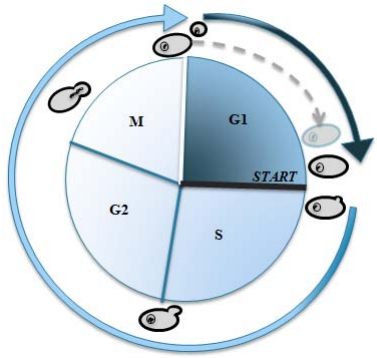

Figure 2.3: Schematic representation of the cell cycle and budding process observed for *S. cerevisiae*.

2.6 Modeling yeast cultivations

Aiming at describing and predicting the behavior of cell cultivations, several mechnistic models of various degrees of complexity have been proposed during the last decades. A very large share of the models proposed for microbial populations is unsegregated [34], i.e. based on an average cell description.

Unsegregated unstructured models are the simplest ones as biomass is considered as a black box: intracellular kinetics are not described, and only the input (e.g. substrate feeding) and output (e.g. substance of interest produced by the microorganism) are accounted for.

Unsegregated structured models form an important class, and incorporate information on the internal mechanism and composition of the microbial mass with the use of several variables, e.g. NADH, precursors, metabolites, ATP, biomass [10]. The number of variables used in such a model should, however, be restricted to a minimum: only variables necessary to obtain information about the most relevant processes of interest should be included. Unsegregated structured models have been used for modeling complex processes, such as yeast intracellular metabolism (e.g. [34]), and morphology-specific growth of filamentous fungi (e.g. [35]). Cybernetic models (e.g. [36, 37]) and different kinds of models based on genomic data (e.g. [38]) belong to this category. These models are beyond the scope of this contribution, and have been reviewed elsewhere [39, 40, 41, 42].

2.7 Describing cell populations: from average descriptions to distributions of cell properties

The work presented in this thesis aims at describing cell populations, rather than considering the cultivation as a collection of cells with identical properties and be-

haviors. In order to achieve this goal, experimental methods for measuring single-cell properties and model frameworks capable of accounting for the observed distributed behaviors are necessary. An overview of the experimental methods and modeling techniques for describing heterogeneous cell populations is provided in the following Chapters 3 and 4, respectively.

Chapter 3

Experimental methods for single cell analysis

Part of this chapter has been included in a review publication:

> Lencastre Fernandes R, Nierychlo M, Lundin L, Pedersen AE, Puentes Tellez PE, Dutta A, Carlquist M, Bolic A, Schäpper D, Brunetti AC, Helmark S, Heins A-L, Jensen AD, Nopens I, Rottwitt K, Szita N, van Elsas JD, Nielsen PH, Martinussen J, Sørensen SJ, Lantz AE, Gernaey KV. *Experimental methods and modeling techniques for description of cell population heterogeneity*. Biotechnol Adv (2011) 29:575-599.

This chapter provides an overview of the motivation behind single cell analysis, as well as an overview of the experimental techniques available for single-cell analysis. Moreover, focus is set, on the one hand, on the use of flow cytometry for studying single-strain microbial cultures, and on the other hand, on the analysis of flow cytometric data.

Microbial populations have traditionally been thought of as large clonal (i.e. isogenic) groups, thus encompassing identical individuals. However, cell populations, particularly when growing in spatially structured environments (like in a large industrial scale bioreactor), can display substantial heterogeneity: individual cells behave differently according to the conditions experienced in their surroundings. This may reflect the different environmental triggers that individual cells experience in a seemingly homogeneous bioreactor [43]. Consequently, the type of environment becomes a key factor for the development of a population and can be considered as a driver of its performance.

3.1 The environment as a driver of microbial population heterogeneity

Generally, environmental stress has been perceived to exert negative effects on microbial populations. In nature, microorganisms are exposed to fluctuating environmental factors, including changes and extremes in temperature, pH, osmolarity, radiation and the concentration of nutrients and toxins. Many microorganisms have developed strategies to cope with such adversities. Central to cell survival are maintaining the integrity of the cell membrane, folding of proteins and the integrity of the DNA [44]. To accomplish this, bacterial cells have, for example, developed systems that sense local conditions, determine when these become deleterious, and stimulate adaptation. Since response levels may be highest under stress conditions, the control networks have been labeled *stress response* systems [45, 46, 47].

The microbial response to stress is generally accomplished by changes in the expression of those genes whose products are required to combat adversity [48]. Some of these stress-induced genes seem to be genuinely specific while others are induced by a wide variety of stresses, and are thus thought to be general stress response genes [45, 49, 50]. Bacteria may use other mechanisms involving physical strategies to survive local adversities, such a sporulation or the use of a flagellum to move to more favorable locations [51]. Microbial responses to stress

lead to cell differentiation and population fragmentation. In such populations, microbial heterogeneity can be observed.

3.2 Experimental methods for characterizing and describing microbial population heterogeneity

As mentioned in the previous section, cell heterogeneity resulting from an environmental pressure implies the co-existence of cells with different physiological states. Being able to characterize and predict the physiological state of individual cells in a microbial population is of great importance in a biotechnological fermentation as 1) the physiological state of the individual cell is the only factor that determines the yield of any product, provided that the required nutrients are present in non-limiting amounts, and 2) consequently, the knowledge of the physiological state is a prerequisite for tuning the fermentation process for optimal performance.

This knowledge has traditionally been acquired indirectly, by measuring a number of parameters like pH, cell density, sugar utilization and product formation. However, as the techniques in the field of molecular biology have improved considerably, the physiological state of cells during the fermentation process has been addressed in much greater detail, primarily by addressing the expression of individual genes, either at the global level by analyzing the transcriptome, or by measuring the expression of genes of particular relevance. Furthermore, the number of studies, based on methods able to quantify properties of single cells, has increased exponentially in the last years [11, 52].

These single-cell level studies generally aim at understanding the mechanisms lying behind the origin of cell heterogeneity, the cause-effect between observed changes in cells and the micro-environmental conditions in the vicinity of these individual cells, as well as the variations in the environment at a macro-scale. Such methodologies may be of physical, chemical and/or molecular nature and

involve a broad range of characteristics, which together give information about the response of populations to environmental cues.

Both microscopy and flow cytometry underwent substantial advances in the last decades, and are nowadays essential tools for monitoring physiological heterogeneity of microbial populations at single cell level. Indeed, a great number of the applications, of both methods, relies on fluorescence monitoring for measuring cellular parameters, such as the case of reporter systems where the cellular component of interest is fluorescent (e.g. reporter proteins such as green fluorescent protein (GFP)). In addition, these methods allow for monitoring other intrinsic cell properties (e.g. cell size), or structural/functional parameters (e.g membrane integrity, DNA content), by applying different staining procedures.

In the frame of this PhD project, only flow cytometry was used for collecting data at single-cell level. Therefore, this chapter will focus on the principles behind flow cytometry, and on analysis of the data collected using his method. For further discussion on the use of microscopy in single-cell studies the reader is referred to the original review publication provided in Appendix Appendix A. Moreover, a recent publication [52] provides a comprehensive review of microfluidics devices and microreactors available for invasive and noninvasive single cell analysis.

3.2.1 Flow Cytometry

Flow cytometry (FCM) is a tool that counts, sorts and examines objects in suspension such as bacteria cells or yeast cells. It is a robust technique that relies on the properties of light scattering, excitation and emission to measure a variety of properties of single cells. The ability of FCM to measure the properties of single cells allows the study of phenotypic diversity of individual microorganisms [53].

When cells pass through a light source, unique electronic and optical param-

eters are measured based on groupings, succession and/or ratios of selected parameters. The information obtained through measurement of the selected parameters can then be linked with different cell properties and components [54].The extensive variety of cellular parameters, that can be studied simultaneously, and the facility to acquire information on how such parameters are distributed in a cell population are the major advantages of flow cytometry as a method for single-cell analysis.

Different methods have been developed in order to study diverse cell properties such as size, intracellular pH and membrane potential that can indicate diverse cellular characteristics such as the levels of cellular components (e.g. DNA, calcium, protein and surface receptors).

Measuring properties of single cells within an entire population can provide a more accurate and descriptive representation of the population than average values attained from traditional techniques [55]. Indeed, due to the possibility of measuring distributed properties in cell cultivations, flow cytometry is a useful tool in the study of heterogeneity in microbial populations [56], and may provide valuable understanding for bioprocess design and control [57]. It has for example been used for monitoring dynamic changes in yeast gene expression [58], for control of biomass concentration [59], for quantification of horizontal gene transfer in bacterial populations [60], as well as for studying heterogeneity of stress gene expression [61]. The role of flow cytometry in molecular biology, with regard to gene reporter systems, has been reviewed by Davey and Winson [62]. Also several different industrial applications of flow cytometry have been reviewed by Díaz et al. [57].

Flow cytometer: principles

As illustrated in Figure 3.1, in a flow cytometer, single cells are hydrodynamically focused in a fluid stream within a carrier fluid known as the sheath fluid. The con-

trol of the pressure in the system and the orifice size allows the establishment of a cellular laminar flow regime. This hydrodynamic focusing of the slower moving sample at the center of a rushed flow stream creates a high speed single cell flow that is intercepted by a light source (usually a laser beam) for the interrogation of particles' properties. The light scattered or emitted by the cells or by cell-associated fluorophores can be isolated and optically separated by collection optics, mirrors and filters. If the angle of deviation of light traveling through the cell is small, it will be detected as forward scatter (FSC). Conversely if the angle of light deviation is big it will be detected as side scatter (SSC). After collection with appropriate wavelength filters, fluorescence is identified on a big angle detector. This light will ultimately trigger a photomultiplier tube (PMT) that augments the signal and finally converts it into a digital signal. In fact, FCM does not usually present real images of the bacteria as fluorescence microscopy, but digital data is produced instead [56, 63, 64, 65].

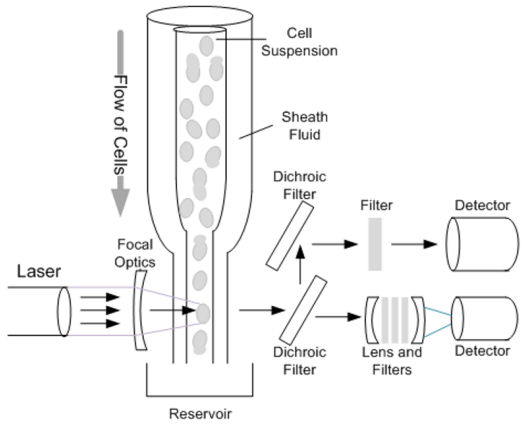

Figure 3.1: Schematic representation of a flow cytometer.

Measurable cell properties

The cell properties measured by FCM can be classified into intrinsic and extrinsic depending on the method of measurement. Intrinsic properties such as size or membrane composition can be studied without the need to label the cells, in the FSC and SSC channels. Moreover, the electronic volume measurement is commonly used for cell size determination [66]. Extrinsic parameters, on the other hand, normally use fluorescent stains or fluorescence labeled probes to study microorganism characteristics and components such as membrane integrity or potential [54]. Given that discrimination of dissimilar cell types and background is attained using fluorescent labels, various fluorescent dyes and labels, also used for microscopy studies, are available for targeting specific biological materials (e.g. nucleic acids) or to signal biological activities (e.g. enzyme activities or membrane potential), leading to a deeper understanding of physiological and metabolic functions in bacteria [67]. Today there is a wide variety of probes and labels available, and the most adequate should be chosen considering the specific strain and cultivation conditions which will be studied [66].

Physiological state of microorganisms

The study of the physiological state of a cell has been redefined by the application of FCM as a tool. The term microbial *viability* gained another dimension upon the disclosure that microorganisms are not just alive or dead, but also have a range of transitional states [68].

Existing fluorescent nucleic acid stains are either able to permeate the membrane or not and the combination of both types of stains, for example SYTO dyes (green fluorescence) and PI (red fluorescence), is used in membrane integrity assays [69, 70, 71]. Furthermore, the cellular membrane potential is used as a test of viability. Depending on the membrane potential and whether the dye used is cationic or anionic, the cell emits (or not) fluorescent signals with different inten-

sities [72, 73].

Measuring the changes in membrane composition directly related with changes in the cellular physiological state is also possible by using the stain 1,6-Dephenyl-1,3,5-hexatriene (DPH) [74]. Testing for enzyme activity is generally done by allowing a membrane permeable nonfluorescent substrate to be taken up by the cell and be metabolized into a preferably impermeable fluorescent substrate (i.e. it cannot cross the cell wall). Based on this principle, respiring and non-respiring cells can be distinguished by the measurement of dehydrogenase activity with, for example, CTC [75, 76] although this method does not always present consistent results [77]. Another example of enzyme activity that can be studied with similar practices is esterase activity by staining with, for example, CFDA [78]. Membrane pump activity can also be assessed by loading different dyes such as PI or ethidium bromide, into the cell and measuring fluorescence reduction [72, 79]. Another test of viability is the measurement of internal pH and its variation with culture change, using a probe whose fluorescent characteristics correlate with changes in pH [80, 81].

Visualization and analysis of flow cytometric data

FCM methods are highly versatile and applicable, and the speed of analysis and the multivariate data sets produced are attractive for developing models for distributed properties. The raw data generated by a flow cytometer when analyzing a sample is a list-mode data file (with the extension *.fcs*), where all the measured properties (FSC, SSC, various fluorescence intensities read by different detectors) are registered for each event (i.e. cell). Commonly used commercial software programs (e.g. DIVA, FlowJo) import .fcs files and provide a user-friendly environment for plotting the data, performing gating analysis as well as making statistical queries.

Histograms, dot-plots (i.e. scatter plots) and contour plots are most often used

to visualize flow cytometric data. In the particular case of dot-plots and contour plots variations can be found. In Figure 3.2, the data for forward scatter and red fluorescence for one single sample (20000 cells were measured) is displayed in different ways. The DNA content in a single cell is quantified based on a red fluorescence intensity, as cells were stained with propidium iodide, after RNA digestion. Specially for two dimensional representations, significant visual differences are observed when comparing dot and contour plots. Contour plots provide a more correct representation of the data [82], as the features of one dimensional distributions (e.g. biomodality) are withheld, and are not always are visible in dot-plots (as in Figure 3.2). These differences may lead to different interpretations [82].

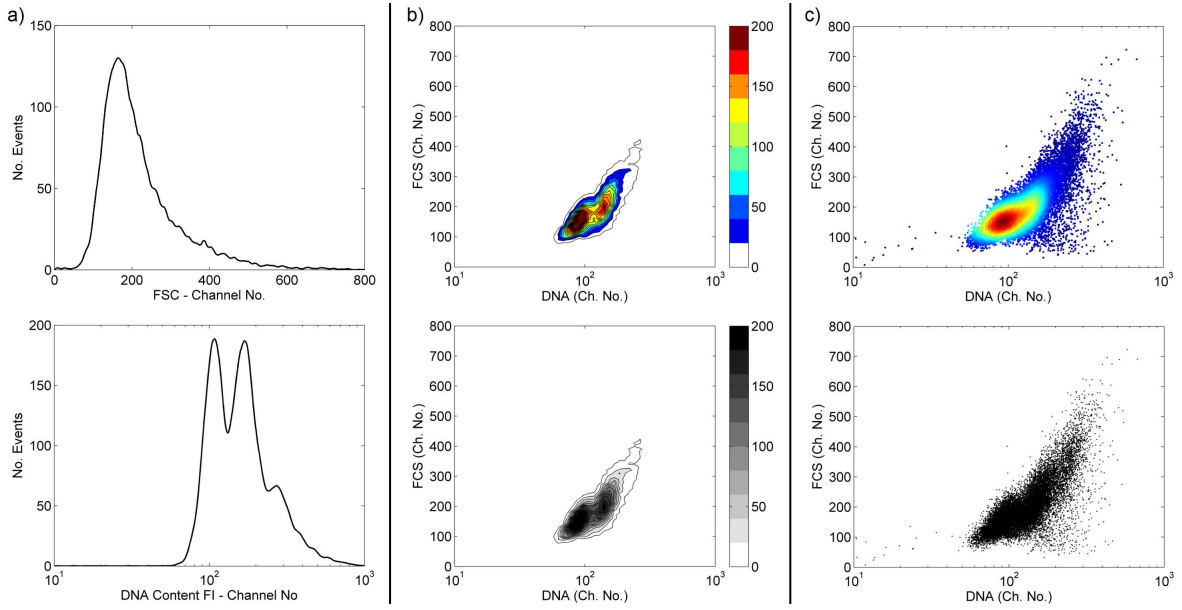

Figure 3.2: Various plots for visualization of flow cytometric data collected for a single sample (20000 events): a) histograms presenting the distribution of FSC and the distribution of red fluorescence intensity, a measure of DNA content; b) bivariate distribution of FSC and DNA content presented as contour plots in color and grey tones; c) bivariate distribution of FSC and DNA content presented as dot plots in color (coding for the density of cells in a given area) and with black markers for each measured event.

Most of the FCM software allows the operator to define areas of specific interest in a procedure called *gating*. Gating works by digitally filtering the FSC, SSC or fluorescence signals, targeting a subset of results. In the case of single-organism samples, gating can discriminate cells of similar properties from a population. Alternatively, gating can be used for identification and counting of different bacterial subpopulations within a complex sample [54, 83, 84].

Practically, gating can be generally defined as a selection of a sub-population by selection of an area (gate) in the one-dimensional histograms or 2-D contour (or dot) plots. Gating strategies consist of sequential steps of sub-population selection and analysis, which allow for the analysis of multivariate data based on visualization in 1 and 2D plots. Traditionally such gates are manually defined (e.g. click and drag on the plot), although other options as the selection of cells within a given contour line (e.g. 95%) are also available in most of the current commercial software.

A simple example of a gating strategy and analysis is presented in Figure 3.3. The aim was to select and analyze the subpopulation presenting lower DNA content. In the first step, the bivariate distribution of FSC and SSC is used to define a gate around the population in order to filter events that may correspond to aggregates of more than one cell (Figure 3.3 a)). Using the bivariate distribution of FSC and DNA content considering only the cells falling within the squared gate imposed in the first step, a second gate is defined (Figure 3.3 b)): in this case a linear gate, defining the subpopulation with low DNA content as the events found to the left of this gate. The FSC histogram for this subpopulation is then compared to the one for the overall population before any gating was performed.

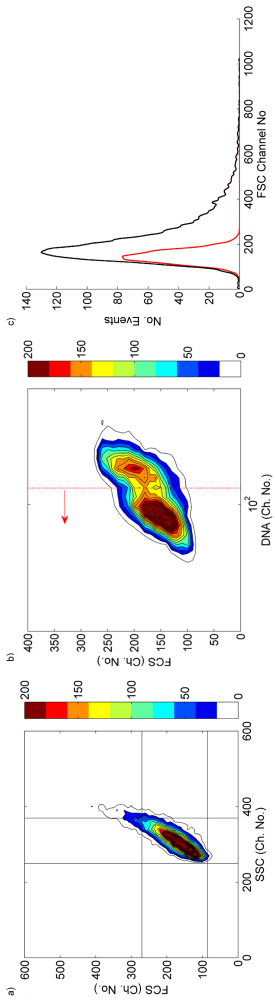

Figure 3.3: Example of gating analysis: a) a squared gate was drawn based on the bivariate distribution of FSC and SSC; b) the bivariate distribution of DNA content and FSC for the cells within the gate in a) is presented and a vertical gate is defined in order to isolate the subpopulation with low DNA content; c) the FSC distribution for the isolated subpopulation is compared to the distribution for the overall population.

More advanced options such as cell cycle analysis or clustering analysis are also available in the commonly used commercial software. All features are, however, designed as black boxes and the user is not offered the possibility of understanding the analysis procedure or make any changes to the algorithms. Furthermore, to my knowledge, these software products do not offer features for tailored data analysis based on user-defined algorithms.

Having in mind that flow cytometry is traditionally used in mammalian cell applications [66], and considering that it is in the field of medical/clinical research that flow cytometry is mostly accepted as an established method [85], caution should be taken when applying black-box advanced analysis methods developed for such larger cells to other systems (e.g. analyzing much smaller microbial cells).

The diversity of research areas where flow cytometry is used as exploded in the last decade (e.g [63, 86, 87, 88]). However, the commercial software programs (designed for traditional clinical applications) do not offer the flexibility necessary to perform more complex analysis for non-standard applications [89]. It is thus not surprising that there is still frequent need for manual analysis of individual samples. This need is, however, indicated as a limiting aspect of the flow cytometry technology [90]. Indeed, due to the subjectivity of manual analysis, independent evaluation of the experimental data and conclusions of a given published study was not always straightforward as it could be desired according to fundamental principles for scientific publication. Moreover, such subjective interpretation makes it generally difficult to apply mathematical models to flow cytometry data.

In order to establish a consensus in the flow cytometry community and broader acceptance of microbial cytometry, a standard defining the minimum information that should be included on scientific publications with regard to procedures, instruments, as well as data analysis and presentation (MIFlowCyt) has been presented by the International Society for Advancements of Cytometry [91, 92]. Concerning data analysis, in particular, MIFlowCyt emphasizes the need for clear gating descriptions.

In the last years, efforts in developing automated gating algorithms have been published (e.g. [90, 93]). Also multivariate and artificial intelligence approaches to flow cytometric data analysis, including cluster analysis and artificial neural networks, have been proposed (as reviewed by [94, 95]). In order to circumvent the limitations of the commercial software in terms of tailor-made analysis, many of these analysis tools have been developed based on the open source statistical language R [96, 97] or the mathematical programming software MatLab [98, 99, 100, 101].

Chapter 4

Modeling Heterogeneous Microbial Populations

This chapter is a modified version of a section in a review article:

> Lencastre Fernandes R, Nierychlo M, Lundin L, Pedersen AE, Puentes Tellez PE, Dutta A, Carlquist M, Bolic A, Schäpper D, Brunetti AC, Helmark S, Heins A-L, Jensen AD, Nopens I, Rottwitt K, Szita N, van Elsas JD, Nielsen PH, Martinussen J, Sørensen SJ, Lantz AE, Gernaey KV. *Experimental methods and modeling techniques for description of cell population heterogeneity*. Biotechnol. Adv. (2011) 29:575-599.

In this chapter, models suitable for describing populations of individual microbial cells are presented and discussed, both in the case of perfectly mixed bioreactors and in the case of large scale reactors where gradients are formed due to mixing limitations.

Similar to chemical systems, the design, control and optimization of bioreac-

tors in many academic and industrial applications have been based on macroscopic first principles models built on balances of extensive properties. However, modeling of bioreactors presents additional challenges as a result of the intrinsic metabolic regulation of the microorganism [102]. Cell variability results in, for example, non-linearities associated with cell growth and division processes even when operated at constant temperature [103]. Further challenges to the formulation of models with an appropriate level of detail and predictive ability arise, thus, when dealing with biological systems.

With the continuous development of the capabilities of techniques available for monitoring of cell properties at single-cell level (see Chapter 3) it is nowadays possible to monitor the distribution of cell properties during a microbial cultivation. In order to take advantage of this knowledge and improve the design and control of bioprocesses, mathematical models able to describe the behavior of a dynamic microbial population are necessary. Such models can support and facilitate the interpretation of the obtained data sets.

4.1 Segregated models: accounting for cell-to-cell variability

Segregated models account for cell-to-cell variation by considering distributed rather than uniform cell properties [13, 104]. They are, thus, necessarily statistical and the degree of structure is at least a scalar such as cell size (mass or volume) or cell age [105].

A more restrictive classification for segregated structured models has been used [106], where structured refers exclusively to a description of a single cell using multiple biochemical substances (i.e. chemically-structured models). In this work, the broader definition (as per [105]) is followed: age and cell mass are descriptors of the cell state, even though they may only indirectly be indicators of the cell metabolism. Therefore, *segregated structured* models (see Figure 4.1) refer to both single variable (e.g. age, mass) models, as well as multivariate,

chemically structured models.

Consequently, *segregated unstructured* models correspond to the cases where the model describes the behavior and size (i.e. total number of cells) of co-existing subpopulations by considering cells by their existence without any further description of the details of cellular metabolism. This would be, for example, the case of activated sludge models that describe the dynamics of heterotrophic and autotrophic biomass subpopulations (e.g. [107, 108]).

In single organism cultivations, subpopulations would, for example, reflect different cell cycle phases. In this case, transitions of cells from one subpopulation to another are possible, and the use of a cell descriptor variable is, therefore, necessary to account for them. This makes it virtually impossible to predict the dynamics of a single organism cultivation using a segregated unstructured model, and further discussion will, thus, focus on segregated structured models.

Segregated structured models can be further classified as single or multi-staged models. The latter account for different cell stages where significant differences in the metabolism are observed, such as budding and non-budding cells [109], daughter and parent cells of different generations [110], different cell cycle phases [104], productive or non-productive phases [111].

Different formulations are possible for segregated structured models. Population Balance Models (PBM) provide the most generic approach to modeling distributed properties, and we focus on this type of models in this chapter. Other simpler models based on ordinary differential equations (ODE) and delay differential equations (DDE) have been reviewed elsewhere [112].

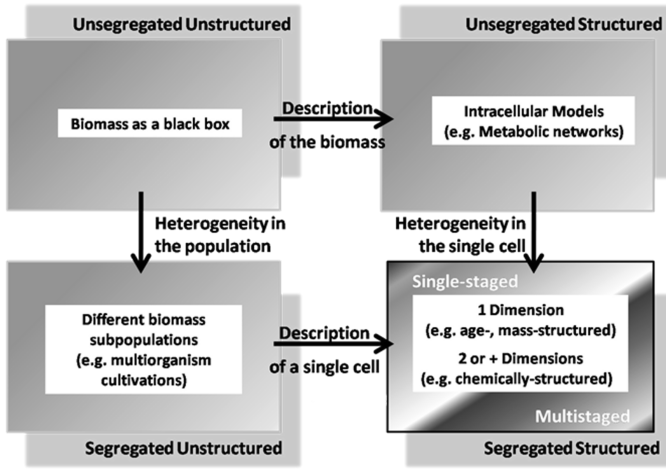

Figure 4.1: Schematic classification of mechanistic models for cell cultivations. Segregated means cell heterogeneity is taken into account, and structured means different cell states are described.

4.2 Population Balance Models for microbial populations

In general, a PBM predicts the temporal change of the cell number distribution, which is characterized by a descriptor variable (e.g. cell age, mass, intracellular metabolites), as result of single cell growth and division into newborn cells. Different formulations are used depending on the cell descriptor variable used.

In Tables 4.1 to 4.4, models for microbial populations that have been published in the last four decades are briefly described. PBM typically consist of a Population Balance Equation (PBE), along with boundary and initial conditions, as well as other coupled equations describing cell division probability and intensity, partitioning of cell content upon division, stage transitions and, in the case of chemically-structured models, cellular kinetics. PBE can be defined as equations of change, i.e. balance equations that account for the various processes that

change the number of cells in a population [104], and take the form of first-order partial integro-differential equations, while the supplementary equations, coupled in a non-linear way, are typically ordinary integro-differential equations [113].

In order to avoid the challenges in solving complex chemically-structured PBM, Mantzaris et al. [113] proposed using a large, but finite number of single cells to represent an entire microbial population. In this Monte Carlo approach, a population of single cells, or cell ensemble, is generated by randomizing kinetic parameters or initial conditions for a single cell model, and it is assumed that the continuity of solutions implicit in the PBM solutions, can be simulated if a large enough number of cells is used. Cell ensemble models have been used to describe respiratory and glycolytic oscillations in yeast populations [114, 115, 116, 117], and further discussions on this approach can be found elsewhere [114, 118].

Table 4.1: One-dimensional Population Balance Models: age as model variable

1-D PBM **Variable: Age**				
Stage	Reactor Mode	Description	Experimental Data	Reference
Single	Continuous	PBM where the environmental conditions are the key cell cycle parameter. Periodic oscillations are sustained by periodic change in the environment, without using specific kinetic expressions. Two models are proposed for binary fission organisms and budding yeast.	Predicted oscillation periods were compared with experimentally observed values, for a range of dilution rates.	[119]
Single	Continuous	PBM was used for predicting periodic behavior of a *S. cerevisiae* cell population and its relation to cell-cycle synchrony.	Experimental observations are taken into account in the formulation of the model.	[120]
Single	Continuous	PBM for synchronous growth of *S. cerevisiae* with asymmetric budding cycle. The model describes sustained oscillations with constant cell number distributions.	Model validation by comparison with experimental data is presented in a subsequent publication.	[121, 122]
Single	Continuous	PBM was used in the design of a controller. Nonlinear feedback control laws are derived in order to attenuate undesired oscillations, or induce synchrony in the *S. cerevisiae* culture.	-	[123]
Active Inactive Dead	Continuous	Model for bioprocess catalyzed by *S. cerevisiae* in a stirred-tank, which is able to reproduce periodic behavior.	Parameter estimation was done using 6 data sets from batch aerated cultivation. Data had been published previously.	[124]
Cell cycle phases Labeled / Unlabeled	Continuous	PBM for human leukemia cells (Jurkat) which models the two subpopulations generated by addition of bromo deoxyuridine.	Age-dependent model parameters were extracted from portioned population data.	[125]
7 development stages from sporangium to mature cel	Batch Fed-Batch	PBM was applied to describe the maturity of sporangium of *Bacillus subtilis* toward the formation of spores. It describes the differentiation phenomenon with associated product formation.	Parameters in the model were determined by fitting the model to experimental data.	[126]

Table 4.2: One-dimensional Population Balance Models: mass as model variable

1-D PBM Variable: Mass				
Stage	Reactor Mode	Description	Experimental Data	Reference
Single	Fed-Batch	Comparison of three simple population models for dynamic simulation of step responses in fed-batch cultures - a simple timer model, a discrete age distribution model which results in a set of ordinary differential equations, and a similar one with discrete mass distribution.	-	[127]
Single	Sequential Batches	In this work, a segregated, structured microbial population balance model is formulated and used to numerically simulate the self-cycling fermentation (SCF) process.	Experimental observations taken into account in the formulation of the model. The model outputs were compared and validated with previously published experimental data.	[128]
Single	Batch	A numerical solution of the mass structured cell PBM in an environment of changing substrate concentration is presented. It can be applied for any type of single-cell growth rate expression, equal or unequal cell partitioning at cell division, and constant or changing substrate concentration.	Experimental observations are taken into account in the formulation of the model.	[129]
Single	Batch startup Continuous	PBM consists of a simple structured description of the extracellular environment, as well it accounts for the three most important metabolic pathways involved in cell growth with glucose substrate of *S. cerevisiae*.	The parameter values were adjusted to achieve qualitative agreement with experimental observations.	[130]
Single	Batch Continuous	The model predicts several situations of batch and continuous growth in which the population density and biomass concentration show opposing trends due to significant variation in the cell mass distribution with time.	-	[131]

Continued on next page

Continued from previous page

Stage	Reactor Mode	Description	Experimental Data	Reference
Single	Continuous	A controller is designed to stabilize steady-state and periodic solutions by regulating the discretized cell number distribution and the substrate concentration. It is based on a dynamic model for the continuous *S. cerevisiae* cultivation.	-	[132]
Daugher Parent Daughter Parent budding	Continuous	Model aims at simulating the effect of dilution rate on the mode of oscillation in continuous cultures of asymmetric budding yeast *S. cerevisiae*.	The growth properties of the yeast were analyzed for continuous cultivations. The distribution of parent and daughter cells in the population was determined microscopically after staining the bud scars and DNA.	[133]
Daugher Parent Non-budding Budding	Continuous	PBM describes the growth of *S. cerevisiae* in spontaneously synchronized continuous cultures.	The structure of the population was identified using oscillating continuous cultures where the division of the cells is synchronized and detectable by large variation of the online measurements (gas exchange rate or heat production rate).	[134]
Daugher Parent Non-budding Budding	Batch	The model framework couples a morphologically-structured representation of the population with population balance theory to formulate a dynamic model for the size distribution of growing yeast populations.	Model validation by comparison with experimental data is presented in a subsequent publication.[135]	[110]
Daugher (Small, Large) Budding	Continuous	The model established a dynamic PBM for asymmetrically dividing yeast. Three special cases are described: step change in growth rate, two transient behaviors following perturbations in the age-distribution.	It is shown how experimental data on transient behavior of a cell population can yield information on single-cell mass-synthesis kinetics and on the manner in which individual cells control certain critical parameters in the cell cycle.	[136]

Continued on next page

Continued from previous page

Stage	Reactor Mode	Description	Experimental Data	Reference
Non-budding Budding	Continuous	The model describes the structural heterogeneity of yeast cell populations (*S. cerevisiae*) and considers the interaction of the population with its environment. Two different situations were investigated: pulse changes of the dilution rate in a continuous process and of the substrate concentration.	-	[109]
Non-budding Budding	Continuous	The model aims at understanding the properties of the microbial biomass in terms of its composition and of the regulation of cell growth and division.	The size at bud emergence and the percentage of budding cells was experimentally determined for a range of dilution rates.	[137]
Cell cycle phases	Batch	A multi-stage population balance model for the growth of ciliated protozoa through its three cell-cycle phases.	Experimental observations are taken into account in the formulation of the model.	[138]
Non-Producing Producing	Continuous	PBM describes the dynamics of cell growth of *S. cerevisiae* during each of the two stages of the cell cycle, including cultivations at limiting substrate and product concentrations.	Experimental observations are taken into account in the formulation of the model.	[111]
Single	Fed-batch	See description above	-	[127]

Table 4.3: Two-dimensional Population Balance Models using mass and age as model variable

2-D PBM **Variable: Age and Mass**				
Stage	Reactor Mode	Description	Experimental Data	Reference
Single	Batch	The model describes the production of ethanol in glucose fermentation of *Zimomonas mobilis*.	Model validation was presented in a second publication [139].	[140]

Continued on next page

Continued from previous page

Stage	Reactor Mode	Description	Experimental Data	Reference
Single	Batch Continuous	The growth-controlled mathematical model of budding yeast predicts theoretical protein and volume distributions.	Compare with protein and volume distributions measured by flow cytometry, for populations growing both in batch and in glucose-limited chemostat cultures.	[141]
Single	Batch	A age and mass structured PBM based on the assumption that only cells from a k^{th} generation originate the $(k+1)^{th}$ generation. This successive generation approach is applied first to one-dimensional model, and then to 2-D one.	Experimental observations are taken into account in the formulation of the model.	[142]

Table 4.4: Multidimensional Population Balance Models

Multi-dimensional PBM **Variable: Physiological state vector**				
Stage	Reactor Mode	Description	Experimental Data	Reference
Single	Batch Continuous	First formulation of a multidimensional PBM: introduction of the concept of physiological state vector.	-	[143]
Single	Continuous	The model aimed at studying the existence of self-similar forms (e.g. time invariant) when each physiological state is scaled with respect to its population average. In this article, each physiological entity was scaled with the respective population average of that entity.	Experimental observations are taken into account in the formulation of the model.	[144]

Continued on next page

Continued from previous page

Stage	Reactor Mode	Description	Experimental Data	Reference
Single	Continuous	A controller is formulated having the PBM as base model. It aims at controlling different moments of the cell mass distribution in a continuous bioreactor by manipulating the dilution rate.	The use of flow cytometry combined with available staining techniques, which allow the on-line measurement of cell property distributions can make the practical implementation of such a control approach possible.	[145]
Multi-staged	Batch	A new and different approach involving randomization of growth rates and compartmentalization is proposed. It aims at circumventing the necessity of having intensity functions for transitions between cell cycle phases, and for which the fission intensity function is state-independent.	-	[146]
Cell cycle phases	Batch	The model is a generalization of the first multidimensional PBM [143], which accounts for passages of cells through a series of recognizable cell cycle phases.	-	[104]

4.2.1 Single variable PBM: mass- and age-structured models

Two different formulations are typically used for a single variable PBM (i.e. one-dimensional or 1-D PBM): age- or mass-structured formulation. The first forms the most simple PBM for microbial cultures as the variation of age with time is unity, and it is, therefore, possible to avoid modeling individual cell growth kinetics [131]. Considering a continuous cultivation in a homogeneous environment, and regarding age as the time elapsed since the birth of the cell, the age-structured PBM expresses the temporal change of the number of cells having a given age.

Although the model is mathematically solvable (e.g. [142, 147]), the adequacy of using age as indicator of the organism state has been questioned [104]. Indeed, a fundamental question arises: is it feasible to monitor the distribution of cell ages, when investigating the development of heterogeneous populations during a dynamic cultivation? In fact, the problem around monitoring of cell age revolves also around the concept of cell age itself. On the one hand, in the case of budding microorganisms, the co-existence of a generation zero of newborn cells, and several generations of mother cells (i.e. of cells which have created one or more daughter cells by budding) could be monitored based on the existence of bud scars and the fact that the cell wall of the newborn cell is synthesized upon budding (e.g. [148]). On the other hand, in the case of microorganisms dividing by fission, it is not possible to distinguish mother and daughter cells. The definition of age is thus intimately connected to the progression through the cell cycle. Extensive work on monitoring cell cycle progression using both microscopy and flow cytometry has been published and reviewed elsewhere [149].

The use of mass as descriptor variable circumvents, at least partly, these problems as distributions of masses can be easily obtained experimentally by using e.g. flow cytometry (see Chapters 3 and 5). Moreover, it has been observed that the size distribution of a cell population responds to changes in the extracellular environment [150]. In this case mass can represent the total cell mass or volume, as well as any conserved property of the cell such as the mass of intracellular components (e.g. total protein, DNA or RNA content).

Although mass distributions are more easily measurable, mass-structured PBM are not able to predict common cellular behaviors such as time lags in the response of cells to an extracellular stimulus [104], and this may explain the sparse use of these models for control, design and optimization of bioprocesses [113]. One solution to this is the use of more than one descriptor variable to better account for cellular metabolism. Multi-stage models can be regarded as an attempt to take more cell descriptors into consideration without increasing the numerical difficulties associated with a two dimensional PBM. Mantzaris et al. [113] proposed a model where different generations (age stages) are considered, as well as (non-)budding sub-stages, and a mass-structured PBM is used to obtain the number distribution for each of the subpopulations. This allowed taking into account the effect of ageing on the growth, budding and division while avoiding the complexity inherent to a two-dimensional (2D) PBM. The definition of transition functions for cell division / birth and budding presents nonetheless an increased effort relatively to single-staged models. Experimental validation of this model is provided in Cipollina et al. [135]. A 2D PBM where a continuum approach is used for both age and mass, has been formulated and numerically solved by Liou et al. [142]. Experimental validation was however not performed. The collection of experimental data for a continuous span of age and mass is, in fact, virtually impossible.

4.2.2 Multivariable PBM: formulation of chemically structured models

Fredrickson et al. [143] proposed the use of a vectorial description of the cell physiological state, which they designated as a physiological state vector. The implementation of such a highly structured PBM has, however, never been achieved, due to the complexity of defining the kernel function for growth, division and partitioning upon cell birth, as well as computational tractability issues that arise when attempting to numerically solve such a model.

4.2.3 Modeling spatial heterogeneity in non-ideally mixed bioreactors

The physiological state of cellular systems and its impact on growth and product formation is the result of a complex interplay between the extracellular environment and the intracellular machinery [113, 151, 152]. As discussed earlier, cells are subjected to spatio-temporal variations in large scale reactors, unlike in laboratory-scale studies under controlled conditions. Indeed, when an individual microorganism circulates through a large scale reactor, it is sequentially exposed to these different local conditions [153]. This may significantly influence the behavior of cellular processes and make conventional (i.e. which assume homogeneous environments) models inapplicable [5].

Coupling fluid models for the extracellular environment with models for the cell population allows for reflecting the interaction between the environment and the physiological state of the cell [5]. A framework suited for capturing the local and global variations in both intra- and extracellular concentrations relies on the link between metabolic network modeling and Computational Fluid Dynamics (CFD) [154]. The use of CFD in modeling bioprocesses is, thus, gaining importance, both in academia and industry [155, 156, 157, 158].

Integration of Computational Fluid Dynamics (CFD)

Computational fluid dynamics has proven to be an efficient and powerful tool for the design and optimization of several flow applications. Typically, a CFD model consists of the three fundamental equations of fluid flow under a given set of conditions to be solved: continuity, momentum and energy equations. Due to the complexity of these equations they are solved numerically to describe the behavior of the system. It is possible to obtain precise predictions of flow and reaction variables using CFD that can be used in scale-up and design applications. Recently, there has been an increased interest in applications of CFD in the (bio)pharmaceutical and biotechnology industries [155, 156, 157, 158]. This

includes analysis of turbulent flow patterns, energy dissipation rates, as well as heat and mass transfer in bioreactors [159], chemical reactions and phase transitions [154] involving several unit operations such as fermentation, mixing and filtration.

Currently, there are two widely used computational approaches for modeling the interaction between phases [160]: the Euler-Euler approach in which different phases are treated mathematically as interpenetrating continua [161, 162], and the Euler-Lagrange approach in which the fluid phase is treated as a continuum whereas the dispersed phase is solved by tracking a large number of particles through the calculated flow field [163, 164, 165].

Bezzo et al. [166] studied xanthan gum production in stirred tanks, and combined the Eulerian approach for the fluid phases with a multizonal model in which the reactor was divided into a limited number of spatial regions. Elqotbi *et al.* [167] implemented an Euler-Euler multi-fluid model to study the interaction of fluid flow, mass transfer and reaction in the fermentation of gluconic acid by *Aspergillus niger* in a gas-liquid stirred fermenter. A constant bubble size was assumed, thus, limiting the possibility of predicting local mass transfer across the phases.

Using the Eulerian-Lagrangian approach, the interaction between the intracellular state of the individual cells of the population and the turbulent flow fields has been studied in a 68 liter [154] and a 900 liter bioreactor [168]. Both structured segregated and unstructured unsegregated approaches were used for modeling the biophase, integrated in a 3D CFD simulation for the reactor. Although Lapin et al. [154, 168] were successful in accounting for the interaction between the individual cells and the spatial concentration gradient caused mainly due to the turbulent flow field, using the stochastic Lagrangian approach - a large number of cells (ca. 10^5) was required to achieve a realistic description of the population, which is computationally intensive.

A dynamic simulation of a heterogeneous cell population in a non-homogeneous

environment can be achieved by coupling CFD and PBM into an integrated framework. The PBM describes the development of the microbial population in a given extracellular environment, while the CFD component allows for the determination of local environmental conditions by calculating flow streams and cell trajectories within the reactor. The integrated CFD-PBM framework should be able to predict local distributed cell properties (e.g. size, composition, age, growth rates, product formation rates) of microorganism populations, while accounting for the changes in the cell physiological state due to different physical (e.g. local gas bubble size, shear stress) and chemical (e.g. local substrate concentration, pH) environments.

As discussed previously, the solution of multivariate PBM requires the use of complex numerical methods. These computational issues are further aggravated with the integration with a CFD model, which would lead to higher calculation times. Nevertheless, it is foreseen that the insight gained by implementing such detailed models will, in the long term, translate into more efficient bioreactor operation [10].

Chapter 5

Populaton dynamics during batch cultivation in an ideally mixed stirred tank reactor

This chapter consists of a slightly extended version of a published research article:

> Lencastre Fernandes R, Carlquist M, Heins A-L, Dutta A, Sørensen SJ, Jensen AD, Nopens I, Eliasson Lantz A, Gernaey KV. 2012. *Cell mass and cell cycle dynamics of an asynchronous budding yeast population: experimental observations, flow cytometry data analysis and multi-scale modeling*. Biotechnol Bioeng (2012) (DOI: 10.1002/bit.24749)

Abstract

Cells in a microbial cultivation present a distribution of phenotypic traits, forming a heterogeneous cell population. A major development in experimental single-cell

studies has taken place in the last decades. It has, however, not been fully accompanied by similar contributions within data analysis and mathematical modeling. Indeed, literature reporting e.g. quantitative analyses of experimental single-cell observations and validation of model predictions for cell property distributions against experimental data is scarce.

This study focuses on the experimental and mathematical description of the dynamics of cell size and cell cycle position distributions, of a population of *S. cerevisiae*, in response to the substrate consumption observed during batch cultivation. The good agreement between the proposed multi-scale model (a PBM coupled to an unstructured model) and experimental data (both the overall physiology and cell size and cell cycle distributions) indicates that a mechanistic model is a suitable tool for describing the microbial population dynamics in a bioreactor. This study therefore contributes towards the understanding of the development of heterogeneous populations during microbial cultivations. More generally, it consists of a step towards a paradigm change in the study and description of cell cultivations, where average cell behaviors observed experimentally now are interpreted as a potential joint result of various co-existing single-cell behaviors, rather than a unique response common to all cells in the cultivation.

5.1 Introduction

As mentioned in Chapter 1, despite traditionally regarded as identical, cells in a microbial cultivation present a distribution of phenotypic traits, forming a heterogeneous cell population. Moreover, the degree of heterogeneity is notably enhanced by changes in micro-environmental conditions. The occurrence of distributions of phenotypic traits such as cell size, enzymatic activities, and growth rate [12], which are often essential for fitness and development of the cells [169], have been observed by using single-cell analysis methods such as flow cytometry or microscopy (see Chapter 3). This type of heterogeneity may originate from stochastic gene transcription, translation and regulation, differences in progression through cell cycle phases, and age distributions due to unequal partitioning

upon division [12, 56].

Frequently, the contribution of different aspects to the development of a heterogeneous population is difficult to distinguish. For example, in the case of *S. cerevisiae*, cell viability after environmental stress (e.g. heat, copper) has shown to be dependent on the cell cycle phase [170, 171], but also to be related to cell age [172]. Robustness to freeze-thaw stress has also been observed to be dependent on the growth conditions prior to the stress [1]. Furthermore, the existence of different microenvironments within a reactor (i.e. spatial heterogeneity) may imply differential responses of cells as they experience a changing extracellular environment in their various trajectories throughout the reactor. Indeed, this fact has been pointed out as the underlying cause explaining differences between cultivations performed in well-mixed lab scale bioreactors and in more poorly mixed large-scale reactors [3].

It is, therefore, in the understanding and description of the interplay between single cell response and the changing environment that the key to build improved predictive models lies. As mentioned in Chapters 3 and 4, the increase in the number of experimental studies at single-cell level was not fully accompanied by similar contributions within data analysis and mathematical modeling. Indeed, literature reporting e.g. quantitative analyses of experimental single-cell observations and validation of model predictions for cell property distributions against experimental data is scarce. This development of predictive models requires detailed experimental observations. Therefore, in order to describe the dynamics of a microbial population in a standardized and quantitative fashion, an adequate set of cell properties has to be selected, and experimental observations have to be conducted and analyzed in a systematic way.

With regard to cell properties, cell size has often been used to describe budding yeast populations under various growth conditions [149]. The choice of cell size relies on its tight coupling to cell growth and division. Indeed, cell size is a key feature affecting cellular design, fitness and function [173], and this is a reflection of the cellular capability of adjusting its growth rate to nutritional availability [174, 175]. The regulation of growth ensures that cells attain a critical size

before initiating the division process [176, 177]. In the particular case of *S. cerevisiae*, two critical sizes corresponding to the regulation points START (committing to budding, or budding transition, see Figure 5.1) and division have been identified (as reviewed in [177]). Hence, using cell size as population physiological state descriptor allows for describing the distribution of cellular states. Experimentally, the distribution of cell size of a population is easily measured by using flow cytometry. In particular, the total protein content has been used as reliable measure of cell size [148].

Cell size distributions of *S. cerevisiae* populations during balanced growth on various limiting substrates, as well as for various dilution rates (i.e. growth rates) have been reported and compared in different studies (e.g. [148, 149, 178, 179]). Larger critical cell sizes (at budding and, consequently, division) have previously been reported for higher growth rates e.g. during exponential growth on glucose relatively to ethanol [135, 148], or with increasing dilution rates in glucose-limited continuous cultivations [176].

Complementary information on the distribution of cells in cell cycle phases can be collected by measuring DNA distributions, yielding a better description of the cellular state [176]. Also age dependency of the critical size upon the budding transition has been evaluated, as time spent in the G1 phase decreases with the number of cell cycles a mother cell has undertaken [179].

As reviewed in Chapter 4, PBM allow for a mathematical description of distributed cell properties within microbial populations [13, 104]. In previously published literature on PBM for microbial populations [104, 129, 138], cell size was used as model variable. Hatzis and Porro [110] proposed a multi-stage PBM accounting for non-budding and budding stages and continuous distributions of cell mass. Additionally, the model distinguished different generations, acknowledging the fact that the critical division size of an individual cell will increase for every cell cycle the cell undergoes. Although the formulation of this model [110] offers the possibility of including the dependence of the critical budding and division sizes on the substrate, this dependence has not been explicitly described neither have simulations under varying substrate conditions been reported. In previous work by

Mantzaris et al. [129], different substrate dependent growth kernels were tested and compared for a cell mass structured PBM. Validation of the assumptions taken or confrontation of model predictions with experimental data was, however, not reported.

The work presented in this chapter aims at understanding and describing the interplay between single cell response, population dynamics and the changing environment, in a systematic way, so experimental observations can be translated into a mathematical model: a PBM describing the cell size distributions for the non-budding and budding populations, during the growth of a *S. cerevisiae*population, in a glucose-limited batch cultivation.

Experimentally, the development of an asynchronous population of *S. cerevisiae* was monitored during the different growth phases, with particular focus on the diauxic shift transition. General trends for cell size (total protein content) and cell cycle phase (DNA) distributions along the cultivation are reported and discussed. Furthermore, a standardized procedure for treatment of the gathered flow cytometric data was established for isolating subpopulations with high content of cells initiating the budding process and preparing for division. The trends identified based on these procedures were used in the definition of the critical budding and division sizes as function of the glucose and ethanol uptake rates: an essential part in the development of a PBM. Additionally, a simple unstructured model was coupled to the PBM in order to evaluate predictions of key changes in the composition of the extracellular environment (i.e. cultivation medium): the consumption of glucose, production and consumption of ethanol, as well as supply and consumption of dissolved oxygen.

5.2 Materials and Methods

The *S. cerevisiae* strain used in this study was the haploid CEN.PK 113-5D (Mat a) with the uracil auxotrophy reversed by transforming a functional URA3 gene.

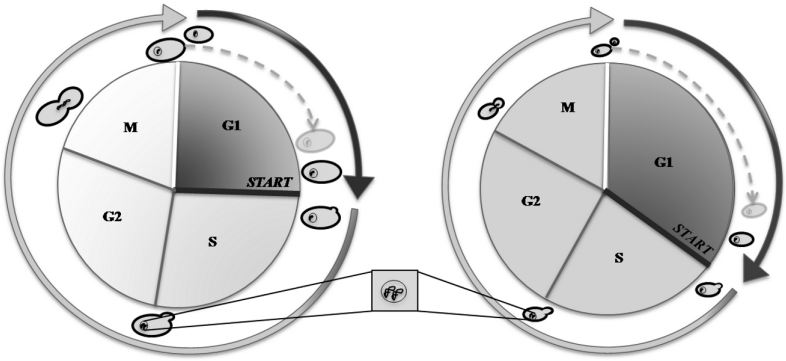

Figure 5.1: Schematic representation of the cell cycle during exponential growth on glucose (to the left) and ethanol (to the right). The dark arrow corresponds to the duration of the non-budding stage (G1 and eventually G0 phases). The START point represents the regulation point that defines the inititation of the DNA replication and budding process, i.e. transition to the S phase. Upon entry to the G2 phase, two copies of the cell DNA are present. Due to their bigger size, the G1 phase is shorter for mother cells than for daughter cells. Generally, the mean cell size is bigger when growing on glucose than on ethanol. The size ratio of bud/daughter cell to mother cell (the two cells originating upon division) is smaller in the case of growth on ethanol.

It was stored in 15% glycerol stocks (liquid medium, -80°C) and plated on YNB-agar plates (6.7 gl^{-1} yeast nitrogen base (Difco, BD Diagnostic Systems, Sparks, MD, USA), 20 gl^{-1} glucose and 20 gl^{-1} agar) and incubated for 2 days at 30°C before use. Inocula were prepared by transferring colonies from plates to Erlenmeyer flasks containing 100 ml defined mineral medium [180] supplemented with 10 gl^{-1} glucose and incubating in a shaking incubator at 150 rpm and 30°C, until mid exponential phase (approx. 10 h). These flasks were used directly for inoculation of the bioreactor (starting $OD_{600} = 0.001$). Braun Biostat 2 liter bioreactors (B. Braun Biotech International, GmbH, Melsungen, Germany) with a working volume of 1.5 l were used for the batch cultivations. The same defined mineral medium supplemented with 5 gl^{-1} glucose was used. The pH and dissolved oxygen tension (DOT) electrode (Mettler Toledo, OH, USA) were calibrated according to standard procedures. Cultivation conditions were set to the following: aeration 1 vvm; temperature 30°C; stirring 600 rpm and pH 5.0 (automatically controlled by addition of 3.0 M KOH). Samples for OD_{600}, high performance liquid chromatography (HPLC) and flow cytometry analysis were taken approximately every

1 hour, or every 30 minutes during the diauxic shift and early growth on ethanol. Samples for OD_{600} were analyzed immediately while samples for HPLC were kept at -20°C and samples for flow cytometry were kept in 15% glycerol at -80°C prior to analysis.

5.2.1 Growth, substrate and products analysis

Growth was monitored by measuring OD_{600} with a UV mini 1240 spectrophotometer (Shimidzu, Kyoto, Japan). The concentrations of glucose, acetate, ethanol, glycerol, pyruvate and succinate were determined by HPLC (Agilent 1100, Agilent Technologies, CA, USA) with a 300mm x7.8mm Aminex HPX-87H ion exchange column (Bio-Rad, Hercules, CA, USA), refractive index detector (RID Agilent 1200, Agilent Technologies, CA, USA) and UV detector set to 210 nm for pyruvate detection (Agilent 1100, Agilent Technologies, CA, USA). The mobile phase was 5 mM H_2SO_4 (aq.), temperature 60° and flow rate of 0.6 $ml\ min^{-1}$. The composition of the outgoing gas was monitored by a 1311 Fast response Triple-gas monitor (Innova Air Technologies, Ballerup, Denmark).

5.2.2 Single-cell analysis

A BD FACSAria III (Becton-Dickinson, Franklin Lakes, NJ, USA) flow cytometer was used for single-cell analysis. For the simultaneous determination of total protein and DNA content, cells were stained with fluorescein isothiocyanate (FITC) and propidium iodide (PI) (SigmaAldrich, Brøndby, Denmark) as described previously [176]. The excitation wavelength for the laser used was 488 nm. Fluorescence emission levels were measured using a band pass filter at 530/30 nm (FITC) and 616/23 nm (PI) . Light scattering and fluorescence levels were standardized using 2.5 μm fluorescent polystyrene beads. 10000 events were recorded, for each sample, with a rate of approximately 1000 events per second.

5.2.3 Computational data treatment

In this work, a systematic approach was applied to the flow cytometry data analysis: standardized procedures were developed for estimation of the critical budding and division sizes based on the experimental total protein and DNA content distributions.

Firstly, a procedure was implemented in order to isolate a subpopulation with a high fraction of cells transitioning from G1 to S-phase, enabling estimation of the critical budding size. It is based on a manual gating strategy described in the literature [176], and relies on isolating the subpopulation presenting an intermediate DNA content (i.e. between 1 copy (1C) and 2 copies (2C) of the chromosome) and thus contained in the interpeak region in the DNA histogram.

A second standardized procedure was developed in order to estimate the critical division size, based on the standard deviation of the 2C peak in the DNA histogram. The critical budding and division DNA bands were defined around the channel number at one standard deviation distance from the peak mode (Figure 5.2). A band width of 10 channel numbers was defined in order to ensure a number of cells in the subpopulation of approximately 500. The critical budding and division cell sizes are defined as the mean total protein content of the cells belonging to the corresponding DNA critical bands. The budding index, i.e. the fraction of cells that have a bud, was estimated using the critical budding DNA band as threshold (Figure 5.3). The robustness of the procedures was assessed by varying the band widths (data not shown). Such variation did not yield a significant effect on the results reported in this work, and the same correlations between critical sizes and substrate availability were observed.

Processing and analysis of flow cytometry raw data was performed by using MatLab® R2009b (The MathWorks, Inc., Natick, MA, USA). The measurement files, exported by the flow cytomer FACSAria II, were imported into MatLab®, using a *fcs data reader* routine (by L.Balkay, University of Debrecen, Hungary), available

on the MatLab® File Exchange website.

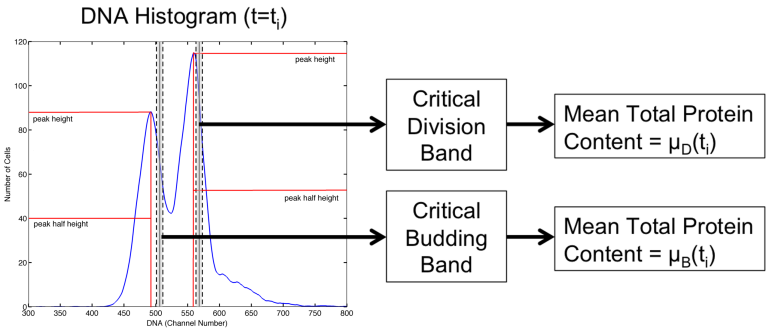

Figure 5.2: Schematic representation of the standardized procedures for definition of the critical budding and division sizes.

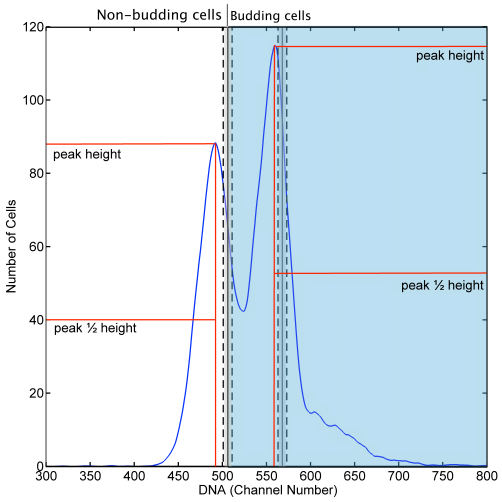

Figure 5.3: Schematic representation of threshold definition used for estimating the budding index (BI, the fraction of the budding cells in the overall population).

5.3 Modeling Aspects

5.3.1 Population Balance Model

A 2-stage PBM using cell total protein content (a measure of cell size) as model variable was developed. Channel number (ch. no.) is used as arbitrary unit for the cell total protein content. Further details on the correlation of channel number to cell size are provided in Appendix B.

The PBM consists of two integro-partial differential equations - the population balance equations (PBE) for the non-budding and budding stages (Equations (5.1) and (5.2)) and the corresponding boundary and initial conditions (Equations (5.3) to (5.5)). $N^{NB}(m,t)dm$ and $N^{B}(m,t)dm$ define the number of cells at time *t* with mass within the interval $[m, m+dm]$, for the non-budding and budding stages respectively. Z describes the extracellular environment (i.e. the concentration of glucose, ethanol and oxygen) at a given time point. A nomenclature list with descriptions and units for the model variables and parameters are provided in Table 5.1.

The left hand side of the PBEs (Equations (5.1) and (5.2)) describes the accumulation of the number of cells in each stage and the growth of cells (i.e. continuous increase of the total cell protein content). In the case of the non-budding stage (Equation (5.1)), the right hand side is composed by a negative budding term, representing cells leaving the non-budding stage by initiating the budding process, and a positive birth term, describing the cells entering this stage as a result of the division of a budding cell into two non-budding cells. In the case of the budding stage (Equation (5.2)), the right hand side is composed of the negative division term (cells leaving the stage as result of division into two new cells) and the positive budding term (upon initiation of the budding process, cells transit

from the non-budding to the budding stage).

$$\begin{aligned}\tfrac{\partial N^{NB}(m,t)}{\partial t} \quad &+\tfrac{\partial}{\partial m}\left[r_m(m,Z)N^{NB}(m,t)\right] = \\ &-\Gamma_B(m\,|Z)\;N^{NB}(m,t) + 2\int_m^{m_f}\Gamma_D(m'\,|Z)P(m,m'\,|Z)N^B(m',t)dm'\end{aligned} \tag{5.1}$$

$$\begin{aligned}\tfrac{\partial N^{B}(m,t)}{\partial t} \quad &+\tfrac{\partial}{\partial m}\left[r_m(m,Z)N^{B}(m,t)\right] = \\ &-\Gamma_D(m\,|Z)\;N^{B}(m,t) + \Gamma_B(m\,|Z)\;N^{NB}(m,t)\end{aligned} \tag{5.2}$$

$$N^{NB}(m_0,t) = N^{NB}(m_f,t) = N^B(m_0,t) = N^B(m_f,t) = 0, m \in [m_0,m_f] \tag{5.3}$$

$$N^{NB}(m,t=0) = \frac{\frac{1}{\sigma_{N^{NB}(m,t=0)}}\phi\left(\frac{m-\mu_{N^{NB}(m,t=0)}}{\sigma_{N^{NB}(m,t=0)}}\right)}{\Phi\left(\frac{mf-\mu_{N^{NB}(m,t=0)}}{\sigma_{N^{NB}(m,t=0)}}\right)-\Phi\left(\frac{m0-\mu_{N^{NB}(m,t=0)}}{\sigma_{N^{NB}(m,t=0)}}\right)} \tag{5.4}$$

$$N^{B}(m,t=0) = \frac{\frac{1}{\sigma_{N^{B}(m,t=0)}}\phi\left(\frac{m-\mu_{N^{B}(m,t=0)}}{\sigma_{N^{B}(m,t=0)}}\right)}{\Phi\left(\frac{mf-\mu_{N^{B}(m,t=0)}}{\sigma_{N^{B}(m,t=0)}}\right)-\Phi\left(\frac{m0-\mu_{N^{B}(m,t=0)}}{\sigma_{N^{B}(m,t=0)}}\right)} \tag{5.5}$$

The boundary condition (Equation (5.3)) assumes that the minimum and maximum cell size (m_0 and m_f) are sufficiently small and big, respectively, so that the number of cells presenting these sizes is zero. In this work, m_0 is defined as 0 ch no. and m_f as 2000 ch. no. Although the experimental data were collected for ch. no. up to 10^4 (see Appendix B), the number of cells observed for channel numbers higher than 2000 was found insignificant, and the discretization upper boundary was set to 2000 ch. no..

The initial distribution for each of the stages is a necessary condition for the PBEs to be solved (Equations (5.4) and (5.5)). To minimize the influence of the inoculum preparation method on the measured distributions, the cell concentration was very low at the beginning of the batch (OD≈0.001). Such low concentrations make the determination of the initial distribution based on a small sample volume prone to error. The distributions measured for the inocula were not considered, as the growth conditions are substantially different from the ones in the bioreactor, upon inoculation. The initial distribution was, thus, not measured, but rather assumed to follow a truncated Gaussian distribution with means equal to $\mu_{N^{NB}(m,t=0)} = 500$ and $\mu_{N^{B}(m,t=0)} = 650$ ch. no. for the non-budding and budding stages, respectively, and a standard deviation ($\sigma_{N^{NB}(m,t=0)}$ and $\sigma_{N^{B}(m,t=0)}$) of 100 ch. no. for both stages. Different mean values were tested without yielding significant impact on the model predictions (data not shown). An increase in the standard deviation has an effect on the model predictions for the initial time points (see Section 5.4).

Growth kernel

The growth rate function, $r_m(m, Z)$ was defined as the product of a mass dependent factor, $k_m m$, and a substrate dependent factor $\lambda(Z)$ (Equation (5.6)). First order kinetics were assumed for the mass dependent factor [110, 129], while the substrate dependent factor $\lambda(Z)$, which can be regarded as a specific growth rate, was derived from the unstructured model for the extracellular environment. The constant k_m operates as a switch that allows for modulating the growth rate

Table 5.1: Description of variables and parameters in the PBM.

Variable Parameter	Parameter Value	Unit	Description
m	-	ch. no.	Cell size
m_0	0	ch. no.	Minimum cell size
m_f	2000	ch. no.	Maximum cell size
m'	-	ch. no.	Cell size of a mother budding cell
Z	-	-	General designation for the extracellular environment
N^{NB}	-	no. cells/ch. no.	Number density for the non-budding stage
N^{B}	-	no. cell/ch. no.	Number density for the budding stage
$\mu_{N^{NB}(m,t=0)}$	500	ch. no.	Mean of the initial number density for the non-budding stage
$\mu_{N^{B}(m,t=0)}$	650	ch. no.	Mean of the initial number density for the budding stage
$\sigma_{N^{NB}(m,t=0)}$	100	ch. no.	Standard deviation of the initial number density for the non-budding stage
$\sigma_{N^{B}(m,t=0)}$	100	ch. no.	Standard deviation of the initial number density for the budding stage
r_m	-	ch. no. /h	Growth rate
k_m	1; Except for stationary phase: 0.4	-	Growth rate constant
$\lambda(Z)$	-	1/h	Specific growth rate (substrate dependent term in the growth rate)
Γ_B	-	ch. no./h	Budding rate
h_B	-	-	Budding probability density function
μ_B	See Table 5.2	ch. no.	Critical budding size (mean of h_B)
σ_B	See Table 5.2	ch. no.	Standard deviation of h_B
k_B	See Table 5.2	ch. no./h	Rate of adjustment of the critical budding size (μ_B)
Γ_D	-	ch. no./h	Division rate
h_D	-	-	Division probability density function
μ_D	See Table 5.2	ch. no.	Critical division size (mean of h_D)
σ_D	See Table 5.2	ch. no.	Standard deviation of h_D
k_D	See Table 5.2	ch. no./h	Rate of adjustment of the critical division size (μ_D)
$P(m, m' \mid Z)$	-	-	Partitioning probability density function

in order to reflect the residual growth observed in the stationary phase. For the

other cultivation phases, k_m is equal to unity.

$$r_m(m, Z) = k_m m \cdot \lambda(Z) \tag{5.6}$$

Budding and division kernels

The budding and division rates, Γ_B and Γ_D (Equation (5.7)), were defined as the product of the growth rate and a hazard function [110, 129]. The latter is based on a density function (h_B or h_D) that describes the probability of a cell of size m to initiate the budding process or to undergo division. Truncated Gaussian probability density functions with mean μ_B or μ_D, respectively, were used. Equation (5.8) shows the budding or division probability density function h_i, where ϕ is a Gaussian probability density function and Φ is a Gaussian cumulative density function. The two mean parameters (μ_B and μ_D) are function of the substrate availability, while the standard deviations (σ_B and σ_D) were assumed to be constant. The numerical values for these parameters are provided in Table 5.2.

Table 5.2: Values for the budding and division parameters.

Unit	μ_B ch. no.	σ_B ch. no.	k_B ch. no./h	μ_D ch. no.	σ_D ch. no.	k_D ch. no./h
Initial ($t = 0$)	$\mu_B^{t=0} = 550$	$0.15\mu_B^{t=0}$	-	$\mu_D^{t=0} = 1100$	$0.15\mu_D^{t=0}$	-
Late growth on glucose $\frac{dG}{dt} < -0.6$ g/lh	$\frac{d\mu_B}{dt} = k_B$	$0.15\mu_B^{t=0}$	-100	$\frac{d\mu_D}{dt} = k_D$	$0.15\mu_D^{t=0}$	-60
Late growth on ethanol $\frac{dE}{dt} < -0.15$ g/lh	$\frac{d\mu_B}{dt} = k_B$	$0.15\mu_B^{t=0}$	-90	$\frac{d\mu_D}{dt} = k_D$	$0.15\mu_D^{t=0}$	-50

$$\Gamma_i(m \mid Z) = r_m(m, Z) \frac{h_i(m|Z)}{1 - \int_{m_0}^{m} h_i(m_k|Z) dm_k} \quad i = \mathrm{B}, \mathrm{D} \tag{5.7}$$

$$h_i(m\,|Z) = \frac{\frac{1}{\sigma_i}\phi\left(\frac{m-\mu_i(Z)}{\sigma_i}\right)}{\Phi\left(\frac{m_f-\mu_i(Z)}{\sigma_i}\right) - \Phi\left(\frac{m_0-\mu_i(Z)}{\sigma_i}\right)} \quad i = \mathrm{B}, \mathrm{D} \tag{5.8}$$

Experimental observations (see Section 5.4) indicated that during late growth on glucose the mean cell size decreased monotonously without a significant variation of the fraction of budding cells in the population (i.e. budding index, BI). The smooth shift of the distribution towards smaller sizes was triggered when a given glucose consumption rate is achieved (see Section 5.4), and modeled by a linear decrease of μ_B and μ_D. The decrease in the critical sizes during the late glucose growth phase is described by Equation (5.9), where I is a switch equal to 1 when the glucose uptake rate is larger than 0.6 $g\, g^{-1} l^{-1}$, and 0 otherwise. A similar behavior was observed during late growth on ethanol (see Section 5.4), but in this case the threshold for ethanol uptake rate is 0.15 $g\, g^{-1} l^{-1}$. The values of rates k_B and k_D (negative values) of both growth phases are provided in Table 5.2.

$$\frac{\partial \mu_i}{\partial t} = k_i \cdot I\left(\frac{dG}{dt} < -0.6\right) \quad i = \mathrm{B}, \mathrm{D} \tag{5.9}$$

At the end of exponential growth on glucose, the transition from the non-budding to the budding stage is arrested when glucose is depleted, and the diauxic shift is initiated. This budding transition arrest was modeled by imposing a budding rate equal to zero. The duration of this arrest is however not dependent on the model variables, and as such to be defined according to the experimental observations: an arrest of 3 hours was defined (reflecting the experimental optical

density curve) in the work here presented.

Birth kernel

The birth term describes the formation of two cells of mass m and $m' - m$ as a result of the division of a budding cell of size m' The birth rate is defined based on the division rate of the mother cell, $\Gamma_D(m'\,|Z\,)$, and the partitioning function $P(m, m'\,|Z\,)$. The latter function describes the ratio of sizes of the daughter cells (originated from the bud) to the mother budding cell, and is defined as a beta probability density function [138] (Equation (5.10)). In this work, the distribution was set as symmetrical ($\alpha = \beta = 50$) during exponential growth on glucose, while during exponential growth on ethanol it was defined as left-skewed ($\alpha = 30, \beta = 60$). The change in the shape parameters reflects the decrease in the ratio of daughter to mother cell size that has been observed experimentally after the diauxic shift [181]. The sensitivity of the model output to the beta distribution shape parameters (α and β) defining the partitioning function is discussed further in this work (see Section 5.4).

$$P\left(m, m'\,|Z)\right) = \frac{1}{B(\alpha,\beta)}\left(\frac{m}{m'}\right)^{\alpha-1}\left(1-\frac{m}{m'}\right)^{\beta-1} \tag{5.10}$$

5.3.2 Unstructured kinetic model for the extracellular environment

An unstructured kinetic model was used to describe how, during a batch cultivation, glucose (carbon source) is consumed, and how ethanol is first produced, and subsequently consumed after glucose is depleted. Also the consumption of dissolved oxygen was modeled as budding yeast oxidizes or reduces glucose de-

pending on the available dissolved oxygen concentration, and its respiratory capacity.

In order to be able to capture the interplay between the cells and composition of the extracellular cultivation environment, a simple kinetic (unstructured) model proposed by Sonnleitner and Käpelli [182] was used for describing the respiratory and fermentative growth of budding yeast.

The unstructured model here used is based on the formulation and assumptions proposed in the original publication [182]. It relies on the stoichiometric equations describing the growth of yeast on glucose and/or ethanol: the purely oxidative growth on glucose (1) or on ethanol (3), or purely reductive growth on glucose (2).

$$C_6H_{12}O_6 + aO_2 + b0.15\,[NH_3] \rightarrow bC_1H_{1.79}O_{0.57}N_{0.15} + cCO_2 + dH_2O \quad (1)$$

$$C_6H_{12}O_6 + g0.15\,[NH_3] \rightarrow gC_1H_{1.79}O_{0.57}N_{0.15} + hCO_2 + iH_2O + jC_2H_6O \quad (2)$$

$$C_2H_6O + kO_2 + l0.15\,[NH_3] \rightarrow lC_1H_{1.79}O_{0.57}N_{0.15} + mCO_2 + nH_2O \quad (3)$$

For each pathway, a set of linear algebraic equations can be formulated when describing the balances for carbon, hydrogen and oxygen. In order for the equation systems to be solvable, one coefficient for each set of equations is to be assumed (based on experimental observations). The yield coefficients (on mass basis) are proportional to the stoichiometric molar coefficients b, g and l, and these yields can be reliabily estimated based on experimental measurements.

$$Y_{XG}^{Oxid} = b\frac{MW(X)}{MW(G)} \quad (5.11)$$

$$Y_{XG}^{Red} = g\frac{MW(X)}{MW(G)} \tag{5.12}$$

$$Y_{XE} = g\frac{MW(X)}{MW(G)} \tag{5.13}$$

For a batch operation mode, the time variation of glucose, ethanol and oxygen is described by Equations (5.14) to (5.16). where G, E, O and X are the concentrations of glucose, ethanol, oxygen, and biomass (in cell dry weight), respectively. $k_L a$ is the mass transfer coefficient, and O^* is the concentration of oxygen in the liquid phase at saturation.

$$\frac{dG}{dt} = -r_{G,\max}\frac{G}{G + K_G}X \tag{5.14}$$

$$\begin{aligned}\frac{dE}{dt} = & \overbrace{r_{G,\max}\frac{G}{G+K_G} - \frac{1}{a}\left(\min\left(r_{O,\max}\frac{O}{O+K_O}, a\cdot r_{G,\max}\frac{G}{G+K_G}\right)\right)}^{q_G^{Red}=q_G^{Total}-q_G^{Oxid}}\cdot jX \\ & -\tfrac{1}{k}\left(\min\left(r_{O,\max}\tfrac{O}{O+K_O} - \min\left(r_{O,\max}\tfrac{O}{O+K_O}, a\cdot r_{G,\max}\tfrac{G}{G+K_G}\right), k\cdot r_{E,\max}\tfrac{E}{E+K_E}\tfrac{K_i}{G+K_i}\right)\right)X\end{aligned} \tag{5.15}$$

$$\begin{aligned}\frac{dO}{dt} = & \; k_L a(O^* - O) - \min\left(r_{O,\max}\tfrac{O}{O+K_O}, a\cdot r_{G,\max}\tfrac{G}{G+K_G}\right)X \\ & -\min\left(r_{O,\max}\tfrac{O}{O+K_O} - \min\left(r_{O,\max}\tfrac{O}{O+K_O}, a\cdot r_{G,\max}\tfrac{G}{G+K_G}\right), k\cdot r_{E,\max}\tfrac{E}{E+K_E}\tfrac{K_i}{G+K_i}\right)X\end{aligned} \tag{5.16}$$

$$\begin{aligned}\lambda(Z) = & \; \tfrac{b}{a}\left(\min\left(r_{O,\max}\tfrac{O}{O+K_O}, a\cdot r_{G,\max}\tfrac{G}{G+K_G}\right)\right) \\ & +g\left(\overbrace{r_{G,\max}\frac{G}{G+K_G} - \frac{1}{a}\left(\min\left(r_{O,\max}\frac{O}{O+K_O}, a\cdot r_{G,\max}\frac{G}{G+K_G}\right)\right)}^{q_G^{Red}=q_G^{Total}-q_G^{Oxid}}\right) \\ & +\tfrac{l}{k}\left(\min\left(r_{O,\max}\tfrac{O}{O+K_O} - \min\left(r_{O,\max}\tfrac{O}{O+K_O}, a\cdot r_{G,\max}\tfrac{G}{G+K_G}\right), k\cdot r_{E,\max}\tfrac{E}{E+K_E}\tfrac{K_i}{G+K_i}\right)\right)\end{aligned} \tag{5.17}$$

In the original model, the biomass accumulation is defined as an autocatalytic reaction. In this work, the corresponding specific growth rate is defined as $\lambda(Z)$ (Equation (5.17)), the extracellular environment factor included in the single cell growth rate (Equation (5.6)) in the PBE. Additionally, the biomass concentration is here predicted using the PBM, and provided as input for the prediction of the extracellular environment variables: glucose, ethanol and oxygen. The overall biomass concentration at any given time point was calculated based on the sum of the zeroth moment (total number of cells) of the cell size distribution for the two stages.

Although it would be expected that cells of small size present a lower dry weight than bigger cells, the error typically associated with the experimental dry weight determination may be quite significant and it would shade differences between smaller and bigger cells. Therefore, the total number of cells was converted into dry weight concentration using a linear regression of cell number to cell dry weight determined experimentally (data not shown), as described by Equation (5.18)

$$X = 4e - 10 \cdot \text{No. Cells} \tag{5.18}$$

The average yield coefficients were estimated by fitting of the model to the experimental data, and are provided in Table 5.3. Although ideally desirable, the parameter estimation was not performed by an optimization routine due to the high collinearity of the parameters in the unstructured model and the complexity of the overall multi-scale model. The design of a parameter estimation routine using a multivariate objective function would be a valuable contribution (e.g. [183]), it however falls out of the scope of this work.

Table 5.3: Description of the variables and parameters in the unstructured model.

Variable Parameter	Parameter Value	Unit	Description
G	-	gl^{-1}	Glucose concentration
E	0	gl^{-1}	Oxygen concentration
O	2000	gl^{-1}	Ethanol concentration
X	-	gl^{-1}	Total biomass (cell dry weight) concentration
DW_{cell}	4e-10	$g\mathrm{cell}^{-1}$	Conversion factor between biomass weight and number of cells
Y_{XG}^{Oxid}	0.4	$gXgG^{-1}$	Yield of biomass on glucose (oxidative metabolism)
Y_{XG}^{Red}	0.06	$gXgG^{-1}$	Yield of biomass on glucose (reductive metabolism)
Y_{XE}	0.4	$gXgE^{-1}$	Yield of biomass on ethanol
$\mu_{E,max}$	0.25	h^{-1}	Maximum specific growth rate for growth on ethanol
$r_{G,max}$	6	$gG\mathrm{gX}^{-1}\mathrm{h}^{-1}$	Maximum specific uptake rate for growth on glucose
$r_{O,max}$	12e-3	$gO\mathrm{gX}^{-1}\mathrm{h}^{-1}$	Maximum specific uptake rate for growth on glucose
K_G	0.5	gl^{-1}	Saturation constant for glucose
K_E	0.5	gl^{-1}	Saturation constant for ethanol
K_O	1e-4	gl^{-1}	Saturation constant for oxygen
K_i	0.1	gl^{-1}	Inhibition constant for glucose
k_La	150	h^{-1}	Mass transfer coefficient

5.3.3 Model implementation and solution

The fixed-pivot technique [184, 185] was applied to a discretization grid in the range $[m_0, m_f] = [0, 2000]$. In this work, an evenly spaced grid of 166 pivots was used. Nonetheless, the same technique could be applied to uneven grids (e.g. geometric) and it has been adapted for a moving grid (the moving pivot technique, [186]).

According to the fixed-pivot technique, the cells generated by division whose mass does not coincide with a pivot are assigned to the neighboring pivots (Figure 5.4). In this case, this reallocation is based on the distance of the new born cell

mass to the neighboring pivots. Similarly to the work of Kumar and Ramkrishna [184] and Nopens [185], the zeroth and first moments (i.e. total number of cells and mean cell size) of the cell size distributions were conserved. The discretized population balance equations for a pivot (in an evenly distributed grid) are given by Equations (5.19) and (5.20), where z corresponds to the boundary of a class where the pivot m is the center point (as illustrated in Figure 5.4).

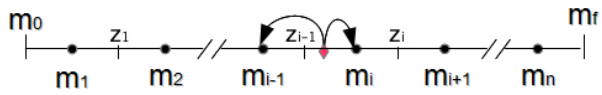

Figure 5.4: Schematic representation of the reallocation of new born cells to the neighboring cell size pivots: black circles mark the grid pivots, the red trapeze marks a new born cell with size between m_{i-1} and m_i (based on [185]).

For notation simplicity, the growth term is designated as θ, the negative division and budding term is represented by γ, and the positive budding and birth terms are noted as ρ. The partial differential PBM equations have been transformed in a system of ODEs for the range of pivots. This ODE system can be summarized in a model matrix (ODE system matrix) as presented in Equation (5.21). By numerical integration of the differential equation system along time the PBM can be solved.

$$\frac{\partial N_{m_i}^{NB}}{\partial t} = \overbrace{\frac{k_m m_i \lambda(Z)}{2(z_i - z_{i-1})} N_{m_{i-1}}^{NB}}^{\theta(m_i)} - \left(\overbrace{\Gamma_B(m_i) + k_m \lambda(Z)}^{\gamma_{NB}(m_i)} \right) N_{m_i}^{NB} - \frac{k_m m_i \lambda(Z)}{2(z_i - z_{i-1})} N_{m_{i+1}}^{NB}$$
$$+2 \sum_{j=i}^{i_{\max}} \left(\underbrace{\left(\int_{m_j}^{m_{j+1}} \frac{m_{j+1} - m}{m_{j+1} - m_j} P(m, m_j) dm + \int_{m_{j-1}}^{m_j} \frac{m - m_{j-1}}{m_j - m_{j-1}} P(m, m_j) dm \right) \Gamma_D(m_j)}_{\rho_{NB}(m_i, m_j)} N_{m_j}^{B} \right) \tag{5.19}$$

$$\frac{\partial N_{m_i}^{B}}{\partial t} = \overbrace{\frac{k_m m_i \lambda(Z)}{2(z_i - z_{i-1})} N_{m_{i-1}}^{B}}^{\theta(m_i)} - \overbrace{(\Gamma_2(m_i) + k_m \lambda(Z))}^{\gamma_{B(m_i)}} - \frac{k_m m_i \lambda(Z)}{2(z_i - z_{i-1})} N_{m_{i+1}}^{B} + \Gamma_1(m_i) N_{m_i}^{NB} \tag{5.20}$$

$$
\begin{bmatrix}
\frac{dN^{NB}_{n_0}}{dt} \\
\vdots \\
\frac{dN^{NB}_{n_{i-1}}}{dt} \\
\frac{dN^{NB}_{n_i}}{dt} \\
\frac{dN^{NB}_{n_{i+1}}}{dt} \\
\vdots \\
\frac{dN^{NB}_{n_p}}{dt} \\
\frac{dN^{B}_{n_0}}{dt} \\
\vdots \\
\frac{dN^{B}_{n_{i-1}}}{dt} \\
\frac{dN^{B}_{n_i}}{dt} \\
\frac{dN^{B}_{n_{i+1}}}{dt} \\
\vdots \\
\frac{dN^{B}_{n_p}}{dt}
\end{bmatrix}
=
\begin{bmatrix}
-\theta(m_0)-\gamma_{NB}(m_0) & -\theta(m_0) & 0 & 0 & 0 & 0 & 0 & \rho_{NB}(m_0,m_0) & \rho_{NB}(m_0,m_1) & \rho_{NB}(m_0,m_{i-1}) & \rho_{NB}(m_0,m_i) & \rho_{NB}(m_0,m_{i+1}) & \rho_{NB}(m_0,m_1) & \rho_{NB}(m_0,m_f) \\
\theta & -\gamma_{NB} & -\theta & 0 & 0 & 0 & 0 & 0 & \rho_{NB}(m_1,m_1) & \rho_{NB}(m_1,m_{i-1}) & \rho_{NB}(m_1,m_i) & \rho_{NB}(m_1,m_{i+1}) & \rho_{NB}(m_1,m_1) & \rho_{NB}(m_1,m_f) \\
0 & \theta(m_{i-1}) & -\gamma_{NB}(m_{i-1}) & -\theta(m_{i-1}) & 0 & 0 & 0 & 0 & 0 & \rho_{NB}(m_{i-1},m_{i-1}) & \rho_{NB}(m_{i-1},m_i) & \rho_{NB}(m_{i-1},m_{i+1}) & \rho_{NB}(m_{i-1},m_1) & \rho_{NB}(m_{i-1},m_f) \\
0 & 0 & \theta(m_i) & -\gamma_{NB}(m_i) & -\theta(m_i) & 0 & 0 & 0 & 0 & 0 & \rho_{NB}(m_i,m_i) & \rho_{NB}(m_i,m_{i+1}) & \rho_{NB}(m_i,m_1) & \rho_{NB}(m_i,m_f) \\
0 & 0 & 0 & \theta(m_{i+1}) & -\gamma_{NB}(m_{i+1}) & -\theta(m_{i+1}) & 0 & 0 & 0 & 0 & 0 & \rho_{NB}(m_{i+1},m_{i+1}) & \rho_{NB}(m_{i+1},m_1) & \rho_{NB}(m_{i+1},m_f) \\
0 & 0 & 0 & 0 & \theta & -\gamma_{NB} & -\theta & 0 & 0 & 0 & 0 & 0 & \rho_{NB}(m_1,m_1) & \rho_{NB}(m_1,m_f) \\
0 & 0 & 0 & 0 & 0 & \theta(m_f) & \theta(m_f)-\gamma_{NB}(m_f) & 0 & 0 & 0 & 0 & 0 & 0 & \rho_{NB}(m_f,m_f) \\
\rho_B(m_0) & 0 & 0 & 0 & 0 & 0 & 0 & -\theta(m_0)-\gamma_B(m_0) & -\theta(m_0) & 0 & 0 & 0 & 0 & 0 \\
0 & \rho_B & 0 & 0 & 0 & 0 & 0 & \theta & -\gamma_B & -\theta & 0 & 0 & 0 & 0 \\
0 & 0 & \rho_B(m_{i-1}) & 0 & 0 & 0 & 0 & 0 & \theta(m_{i-1}) & -\gamma_B(m_{i-1}) & -\theta(m_{i-1}) & 0 & 0 & 0 \\
0 & 0 & 0 & \rho_B(m_i) & 0 & 0 & 0 & 0 & 0 & \theta(m_i) & -\gamma_B(m_i) & -\theta(m_i) & 0 & 0 \\
0 & 0 & 0 & 0 & \rho_B(m_{i-1}) & 0 & 0 & 0 & 0 & 0 & \theta(m_{i-1}) & -\gamma_B(m_{i-1}) & -\theta(m_{i-1}) & 0 \\
0 & 0 & 0 & 0 & 0 & \rho_B & 0 & 0 & 0 & 0 & 0 & \theta & -\gamma_B & -\theta \\
0 & 0 & 0 & 0 & 0 & 0 & \rho_B(m_f) & 0 & 0 & 0 & 0 & 0 & \theta(m_f) & \theta(m_f)-\gamma_B(m_f)
\end{bmatrix}
\begin{bmatrix}
N^{NB}_{n_0} \\
\vdots \\
N^{NB}_{n_{i-1}} \\
N^{NB}_{n_i} \\
N^{NB}_{n_{i+1}} \\
\vdots \\
N^{NB}_{n_p} \\
N^{B}_{n_0} \\
\vdots \\
N^{B}_{n_{i-1}} \\
N^{B}_{n_i} \\
N^{B}_{n_{i+1}} \\
\vdots \\
N^{B}_{n_p}
\end{bmatrix}
\tag{5.21}
$$

Coupling of the PBM to the unstructured model

As previously described, the unstructured model consists of three ODEs describing the variation over time of concentrations of glucose, ethanol and oxygen. An iterative scheme was used to solve the PBM and the unstructured kinetic model. A schematic representation of the model solution steps is presented in Figure 5.5.

Time steps of 0.1 h were defined, ensuring that the variation in the concentrations of glucose and ethanol per time step was smaller than 0.1 gl^{-1} - the approximate experimental error in the determination of these concentrations (see Section 5.2). Smaller time steps did not yield significant differences in the model predictions (data not shown), confirming solution convergence. All model simulations were performed using MatLab ® R2009b (The MathWorks, Inc., Natick, MA, USA). The ODE solver ode15s was used for solving the PBM ode system, whereas the unstructured model ODE system was solved using solver ode23 (absolute and relative tolerance were set to $1e-3$, while default values were taken for all other solver options). The total number of cells was converted into dry weight concentration using a linear relation, as previously discussed and defined by Equation (5.18).

For each iteration , a biomass concentration vector corresponding to 10 time instants in the time span of the iteration is provided as input to the ODE solver for integration of the unstructured model equations (Equations (5.14) to (5.16)) for the same time span of the iteration. The updated concentrations of glucose, ethanol and oxygen are used to re-calculate the substrate dependent factor $\lambda(Z)$ in the growth rate, and eventually update the budding and division parameters, μ_B and μ_D, that depend on the substrate consumption rates.

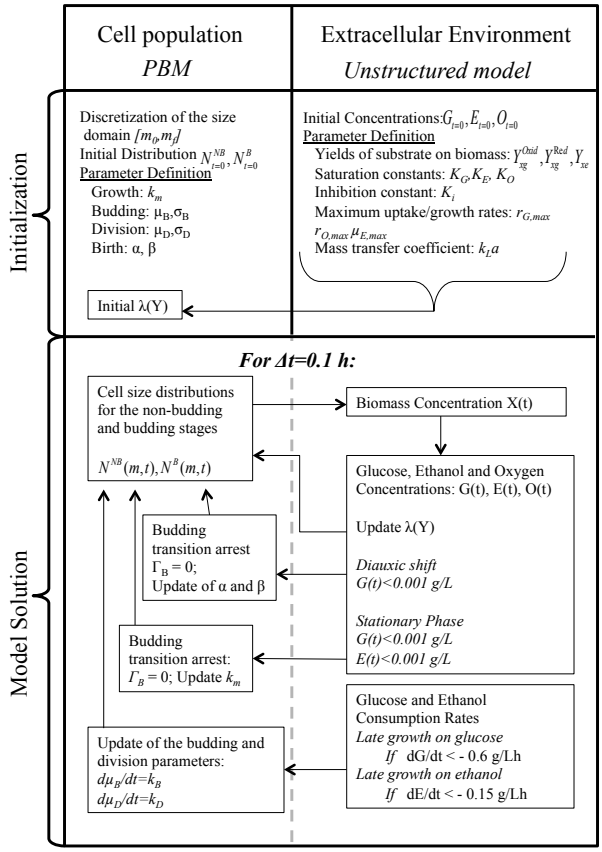

Figure 5.5: Iterative procedure used for solving the two-stage PBM and unstructured model for a batch cultivation.

5.4 Results and Discussion

Triplicate glucose limited aerobic batch cultivations of *S. cerevisiae* were performed. The three cultivations presented similar growth and production profiles (see Appendix B). The respective biomass (optical density), glucose and ethanol concentration curves followed the typical patterns for aerobic yeast cultivations. The population proliferated exponentially first on glucose with a specific growth rate of 0.41$\pm$0.001 h^{-1}, and of 0.10$\pm$0.02 h^{-1} during the exponential growth on ethanol. The distributions of cell total protein content (a measure of cell size) and DNA were monitored, by flow cytometry, during the four cultivation phases. Generally, there was a good agreement between the three replicate cultivations with regards to the flow cytometry measurements (see Appendix B).

The predictions of the integrated model describe both the variation of macroscopic variables - glucose, ethanol and dissolved oxygen, as well as the overall biomass concentration - and the distributions of single-cell total protein content for the non-budding and budding populations during glucose-limited batch cultivation. With regard to the macroscopic variables, the comparison of the model predictions and experimental results (Figure 5.6) shows a good agreement: the model is able to describe the growth phases (on glucose and ethanol) as well as the intermediate diauxic shift phase. Based on the increase of number of cells predicted by the model, the model predicted specific growth rate was 0.44 h^{-1}. This corresponds to a deviation of approx. 7% between the estimation and experimental data.

5.4.1 Mean Total Protein Content

The mean total protein content of the population decreased throughout the cultivation (Figure 5.7). During exponential growth on glucose, the population showed a higher mean total protein content relative to the other cultivation phases (i.e. larger cell size), corresponding to a channel number of approximately 700. In late

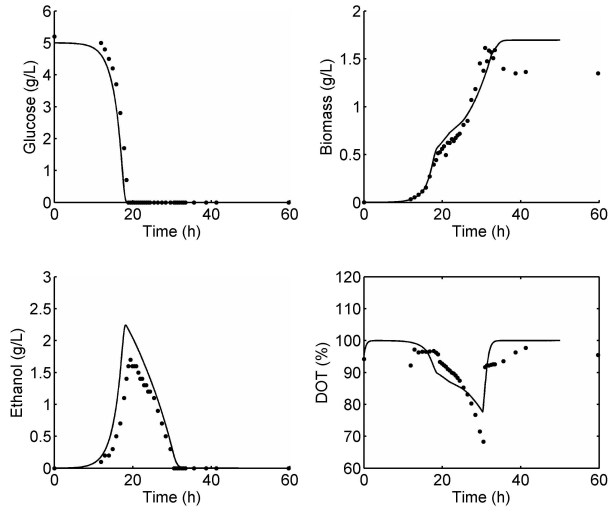

Figure 5.6: Variation of glucose, overall biomass, and ethanol concentrations and dissolved oxygen tension along the cultivation: model predictions (full line) and experimental observations (full circles).

exponential phase during glucose assimilation, the mean cell size decreased to a ch. no. of approximately 450. An additional decrease of the mean cell size to a ch. no. of 350 was observed upon the depletion of glucose. Similarly to what was observed during growth on glucose, the mean cell size decreased during late exponential growth on ethanol, from a mean ch. no. of approximately 300 to 200. The mean cell size remained approximately constant during stationary phase until t =60 h, when the cultivations were terminated. Also the predicted mean total protein content for the overall population was in good agreement with the experimental observations.

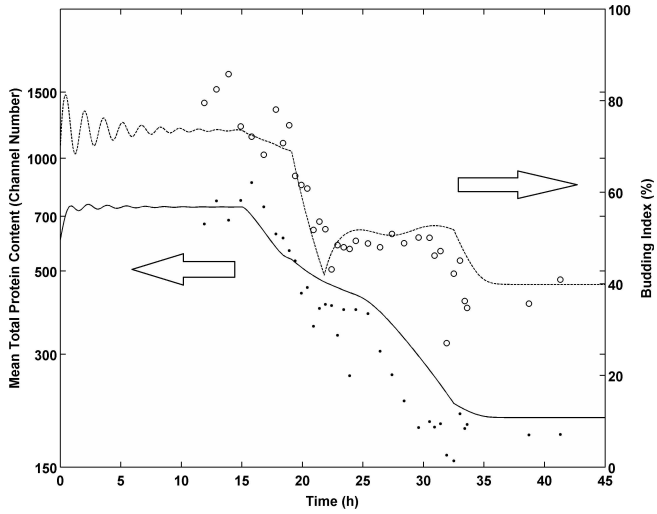

Figure 5.7: Comparison of model predictions for the mean cell size or total protein content (full line) and budding index (dashed line) to experimental observations (cell size as dots and budding index as open circles).

5.4.2 Budding index

The variation in the distribution of cells through the cell cycle phases can be seen even clearer when observing the variation of the budding index (BI), along the cultivation (Figure 5.7). Despite deviations between the estimations based on flow cytometry DNA distributions (as here reported) and microscopic counting have been observed (see Appendix B), the BIs estimated from flow cytometric data for time points during exponential growth on glucose and ethanol, presented in this work, are in good agreement with values reported in the literature. A BI of approximately 80% was observed during growth on glucose, which is in agreement with reported BI values for cultivations with similar specific growth rates [187, 188].

The initial oscillations of the budding index and subsequent damping can be interpreted as a desynchronization of the population resulting from the fact that the critical sizes are described by a probability distribution function rather than a unique discrete value [129, 142]. An increase in the standard deviation assumed for the initial distribution results in a quicker damping of these oscillations.

The abrupt decrease of the BI to approximately 48% from about t =18 h onwards coincides with the depletion of glucose. A similar BI has been observed during exponential growth in an ethanol-limited batch cultivation [188]. During stationary phase, an accumulation of G1 cells took place analogously to the observations during the diauxic shift, resulting in a BI of approximately 42% after 60 hours of cultivation.

Generally, the model predictions are in good agreement with the estimations based on the experimental DNA distributions here reported. Indeed, the accumulation of smaller non-budding cells observed in the diauxic shift is successfully described by the model, and reflected in the continued decrease of the mean cell size and accentuated the decrease of the budding index during this phase of metabolic rearrangement (Figure 5.7).

When ethanol is depleted and cells enter the stationary phase, the typically observed accumulation of non-budding cells takes place at slower pace than the one observed for the diauxic shift. Therefore, an arrest of the budding transition is imposed as well as decreasing the constant k_m from unity to an assumed value of 0.4. Notwithstanding the capability of the model of describing the decrease in the mean cell size and budding index, the transition into stationary phase is rather complex, and not yet fully understood [189]. Therefore, the simplistic model description presented (by decreasing k_m and arresting the budding transition) should not be expected to describe such complexity. More research is required to fully unravel this mechanism.

5.4.3 Cell size and cell cycle position distributions

In Figure 5.8, the bivariate distributions of total protein and DNA content measured along the cultivation are presented as contour plots. The distributions of each cell property (in the shape of histograms), for all sample times, are provided in Appendix B. From the bivariate distribution, the significant changes in the structure of the population can be easily visualized. During the growth on glucose, a large part of the population consisted of bigger cells presenting 2 copies of the DNA (i.e. cell in the G2+M cell cycle phases). As reflected in the BI profile (Figure 5.7), during the diauxic shift, an accumulation of cells with lower DNA content (corresponding to cells in the G1 cell cycle phase) and smaller size was observed, resulting in a clearer cloud on the bottom left corner of the bivariate distribution (Figure 5.8, t =20.4 h). During growth on ethanol, a more even distribution of the cells containing 1 or 2 copies of DNA, in comparison to growth on glucose, was observed. This is illustrated in Figure 5.8: while for exponential growth on glucose (t =16.8 h) only one high density cloud (red colored) is observed, two clouds with approximately the same densities (blue colored) are observed during the growth on ethanol (t =23.9 h).

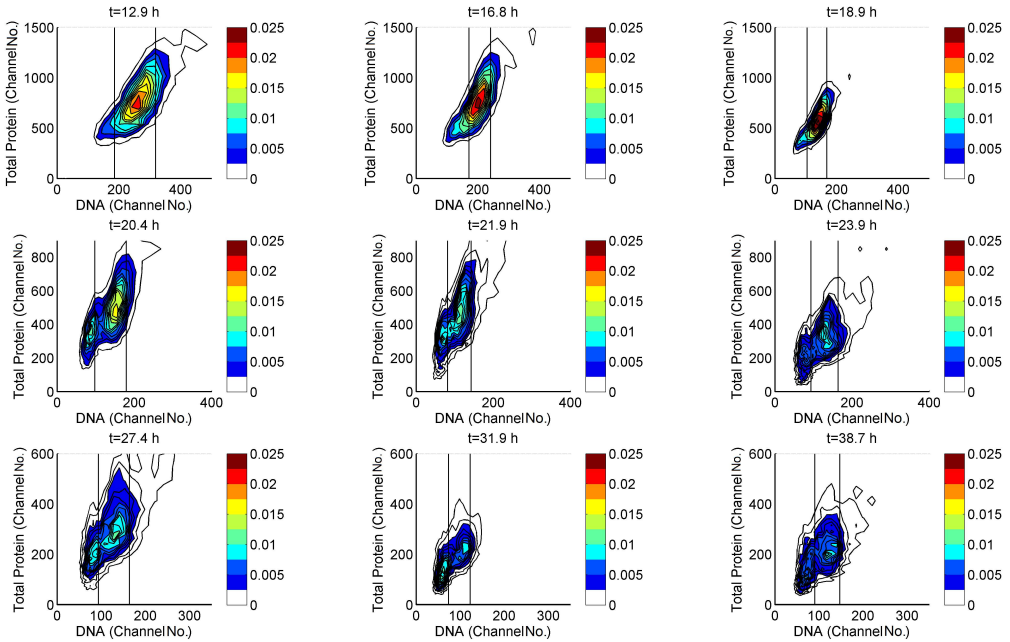

Figure 5.8: Bivariate distribution of total protein content and DNA during a batch cultivation of *S. cerevisiae*: exponential growth on glucose (t =12.9, 16.8 h), diauxic shift (t =18.9, 20.4 h), exponential growth on ethanol (t =21.9, 23.9, 27.4 h), and stationary state (t =31.9, 38.7 h). The color code corresponds to the number density of the cells. The vertical lines correspond to the critical budding (to the left) and division (to the right) threshold identified based on the DNA distribution.

In Figure 5.9, the total protein content distributions predicted by the model are compared to the experimental ones. The model predictions successfully describe the general shift of the distributions towards smaller sizes. The biggest difference between model predictions and experimental distributions is observed during the diauxic shift. While the model shows clear bimodal distributions around t =21 h, the experimental distributions are unimodal. The experimental distributions for each of the non-budding and budding subpopulations may present larger variances than described in the model, and this may explain the loss in separation of the peaks corresponding to non-budding and budding cells (unimodality) observed experimentally. Additionally, the model does not acknowledge the existence of different generations whose critical sizes increase with the generation age [110], leading to intermediate subpopulations.

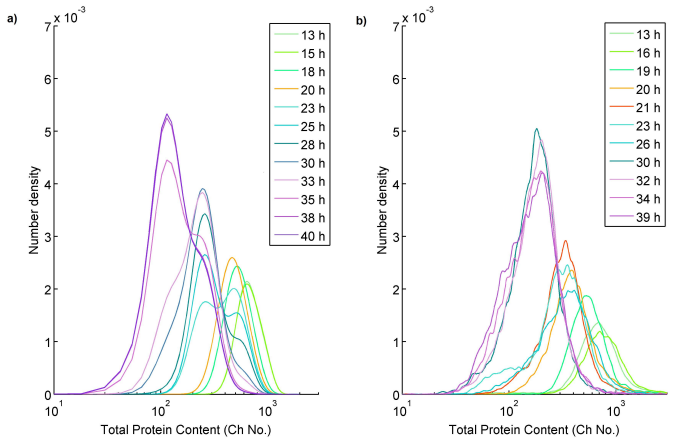

Figure 5.9: Total protein content distributions for the overall population: a) model predictions b) experimental observations. The color code reflects the different cultivation phases: exponential growth on glucose (green), diauxic shift (orange-red), exponential growth on glucose (blue), stationary phase (violet).

5.4.4 Critical budding and division sizes: dependence on substrate availability and uptake

The model presented in this work relies on the fact that the cell size regulation has been identified as a key aspect in the growth regulation in response to changing environmental conditions [177, 190]. Understanding the dependence of the critical budding and division sizes on the extracellular environment is essential for understanding the population dynamics and for future modeling of the dynamics of size and cell cycle position distributions. Indeed, to define the critical budding and division sizes as a function of the substrate availability is essential for the definition of a dynamic population model (see Section 5.3).

In order to assess how the mean cell size of a subpopulation with a high content of cells initiating the budding process, i.e. cells in the critical budding size range, changed during the batch cultivation a standardized procedure was developed and applied to all samples (see Section 5.2). Not surprisingly the critical budding size followed a similar evolution as the overall mean cell size (Figure 5.10), and it was constant during early growth phases and decreased in the later growth phases and diauxic shift. Similarly, a standardized procedure was applied to estimate the critical division size at each sample time point. Also the critical division size accompanied the general shift of the overall population towards smaller sizes along the cultivation (Figure 5.10).

In this work, particular attention was paid to understand the population dynamics observed during the transition between growth on glucose and ethanol. From the variation of the estimated critical budding and division sizes along the cultivation during the late growth phases on glucose or ethanol (Figure 5.9), cells seem to adjust to the decreasing substrate availability by a smooth shift of the cell size distributions towards smaller sizes, and a slight decrease in the BI (Figure 5.7). Contrarily, an abrupt change in the cell cycle position (sudden sharp decrease of the BI) was observed upon glucose depletion and beginning of the diauxic shift. These observations are in agreement with the two different mechanisms proposed by Brauer et al. [187] when explaining the changes in the gene

expression patterns observed during (i) the late growth phases - *continuous homeostatic metabolic adjustment* - and (ii) diauxic shift - *discontinuous metabolic remodeling*.

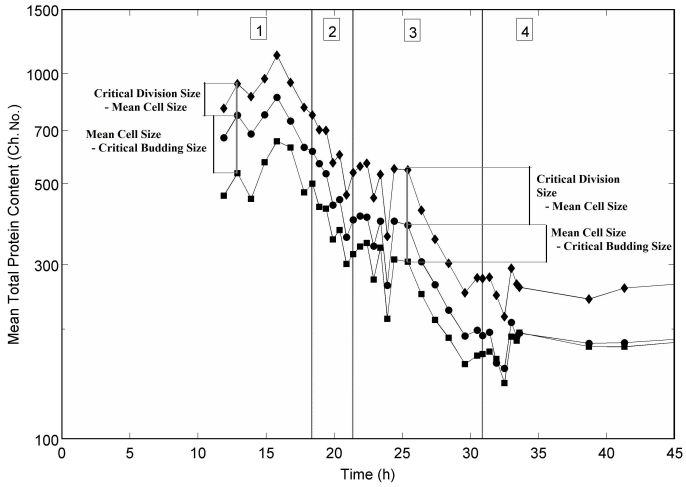

Figure 5.10: Variation of the mean total protein content (i.e. cell size) for the overall population (circles), the critical budding size (squares) and critical division size (diamonds). The vertical lines define the four phases of the cultivation: (1) exponential growth on glucose (2) diauxic shift (3) exponential growth on ethanol (4) stationary phase. The distances between cell size and the critical budding and division sizes reflect the cell cycle distribution (fractions of non-budding and budding cells within the population).

Cell adjustment during late growth phases

As mentioned above, during the late growth phases, the cell size distribution smoothly shifted towards smaller sizes, while the BI generally remained constant (Figure 5.7). In this work, the initial glucose concentration was 5 gl^{-1}, and the decrease in cell size was observed approximately 2 hours before the diauxic shift occurred, corresponding to a glucose concentration above 3 gl^{-1}. Brauer et al. [187] performed cultivations with an initial glucose concentration of 2.5 gl^{-1}, and

significant changes in gene expression were observed when the glucose concentration was approximately 0.5 gl^{-1}. It thus seems unlikely that the onset of the metabolic adjustment is triggered by a certain concentration of glucose in the medium. Interestingly, when comparing the glucose consumption and biomass growth curves of Brauer et al. [187] with those presented in Figure 5.6, the start of the adjustment in gene expression and decrease in the mean cell size coincided with a steeper increase in biomass concentration (OD) and consumption of glucose. The glucose consumption rate estimated for this time point was approximately 0.6 $gl^{-1}h^{-1}$. A similar rate was estimated from the publication of Brauer et al. [187]. With regard to the growth phase on ethanol, the trend of cell size distribution and the critical budding size was similar to the behavior during growth on glucose. A similar adjustment behavior triggered by threshold ethanol consumption rate was thus assumed. This threshold was estimated to be approximately 0.15 $gl^{-1}h^{-1}$.

The dependence of the critical budding size on the specific growth rate has been reported elsewhere [176], and is implicitly related with the uptake rates [191]. The link between the critical division size and the uptake rate is thus not unexpected. In fact, Youk and Oudenaarden [191] discuss the interaction between glucose perception and import (uptake) on the growth: for wild-type *S. cerevisiae*, when glucose is not too low, the effect of glucose perception disappears and the uptake perception is dominant. A more in depth validation of these assumptions and numerical values for the threshold uptake rates here proposed could for example be achieved by comparison to the behavior of strains with improved glucose or ethanol affinities. Such validation lies, however, outside the scope of this work.

Cell rearrangement upon diauxic shift

During the diauxic shift, an abrupt decrease of BI was observed (Figure 5.7) together with the appearance of a subpopulation of smaller cells (Figure 5.8). The cell rearrangement is likely to have more underlying reasons than merely a modulation of the critical budding and division sizes. In fact, it has been reported that

the difference in cell size between the two cells originating upon division - the bigger mother cell and the smaller new born cell - is more accentuated for growth on ethanol than for growth on glucose [192, 193].

Furthermore, the diauxic shift is typically characterized by arrest in the increase of biomass. As reviewed by Alberghina et al. [181], the budding transition is likely to be resumed before the division, leading to a small increase of the BI after a sharp downward shift. This can be described by an abrupt change in the partition shape parameters. A small downward peak and consequent recovery were observed in Figure 5.7 (t =22.9 h); however it should not be considered as conclusive given the nature of the BI estimation procedure. Such an abrupt change of the partition shape parameters allowing smaller cells to eventually initiate DNA replication and the budding process were however not captured when considering the mean critical budding size (Figure 5.10). Indeed, the standardized procedure developed in this work based on analyzing the DNA histograms is not able to capture abrupt dynamics: it is necessary that a sufficient fraction of cells with a smaller size initiating DNA replication is present in the population, for its effect on the mean size of the subpopulation isolated as the critical budding band to be observed. The sensitivity of the method may be improved by increasing the number of analyzed cells for the cases when faster dynamics are expected. Furthermore, cells of different ages could be distinguished experimentally by additional staining [171], allowing specifically for isolating and comparison of the newborn cells originated in the different growth phases.

5.4.5 Sensitivity of the model output to the partition function parameters

The description of the partition function as a beta distribution (Equation (5.10)) implies that the two shape parameters, α and β, are defined. In a previous theoretical study by Mantzaris et al. [111], a symmetrical distribution with shape parameters α and β equal to 40 was assumed without experimental evidences. Other studies [110, 194, 195, 196] assumed the same symmetrical distribution

and shape parameter values.

In this work, the numerical values of each of the shape parameters have also not been directly estimated from the experimental data, but the effect of these parameters on the model outputs has been assessed. An experimental estimation of the partition distribution (and corresponding shape parameters) implies determining the cell size distribution of a subpopulation of newly born cells: not only the smaller cells originating from the bud but also the bigger cells that produced the bud. This could be done taking a similar approach to Porro and Srienc [178]: at a given time point all cells in a sample were stained with ConA-FITC (labeling the cell surface) and resuspended in growth media. Assuming that the first partly stained cells to appear correspond to cells that have just divided, the total protein content of this subpopulation can be determined: a symmetric distribution will indicate a symmetric partitioning (more equally sized cells) while an asymmetric (eventually bimodal) distribution would indicate an unequal partitioning. It is, however, not clear if this experimental procedure would allow for a quantitative estimation of the shape parameters.

The numerical values chosen - a symmetrical distribution initially ($\alpha = \beta = 50$) and a left-skewed distribution from the diauxic shift onwards ($\alpha = 30$ and $\beta = 60$) - correspond to the best fit of the model predictions to the experimental data. The change in the parameters from one growth phase to the next, reflects the fact that the two cells originated upon division are more similar in size during growth on glucose than during growth on ethanol.

Model simulations for different parameter combinations were compared in order to assess the sensitivity of the model to the two shape parameters. Firstly, simulations with different combinations of values for α and β during growth of glucose were carried out, while maintaining those after the diauxic shift at $\alpha = 30$ and $\beta = 60$ (Case I). The resulting model predictions are compared in Figure 5.11. Secondly, simulations with different combinations of values for α and β after the diauxic shift were made, while maintaining the initial values $\alpha = \beta = 50$ for growth on glucose constant (Case II). The resulting model predictions are compared in Figure 5.12. The beta distributions corresponding to the various parameter com-

binations are illustrated in Appendix B.

Concerning the sensitivity of the model to the partitioning function parameters provided initially (Case I, Figure 5.11), the more asymmetrical the beta distribution (i.e. bigger difference in the size of the two cells resulting from division) the slower the growth is. The slower growth leads to a slower consumption of glucose and a consequent time delay in the ethanol and dissolved oxygen profiles (Figure 5.11 a)). This can be explained by the existence of smaller cells (lower overall mean cell size - Figure 5.11 b)) which require a longer time to go through the cell cycle, i.e. spend more time in the non-budding stage (lower budding index - Figure 5.11 c)). Furthermore, the existence of smaller non-budding cells results in greater oscillations of the budding index (Figure 5.11 c)).

As expected, a similar effect is observed when different parameter combinations are used for redefining the partitioning function upon the diauxic shift (Case II). The sensitivity of the ethanol and dissolved oxygen (model outputs) is however lower in this case (Figure 5.12 a)). After the abrupt decrease caused by the arrest in the budding transition, the budding index increases upon resuming the transition: the more symmetrical the partition function, the higher the predicted budding index is upon resuming the budding transition. This is not surprising when considering that, for a more symmetrical distribution, the cells accumulating in the non-budding stage (during the arrest of the budding transition) have more similar sizes and, upon resuming the budding transition, will become budding cells at approximately the same time causing a larger increase in the budding index.

5.5 Conclusions

This study focuses on understanding and modeling the dynamics of a yeast population in terms of development of the cell size and cell cycle distributions along a batch cultivation. It consists of an example of a quantitative integrated analysis of general physiology data (i.e. substrate and metabolite concentrations, OD)

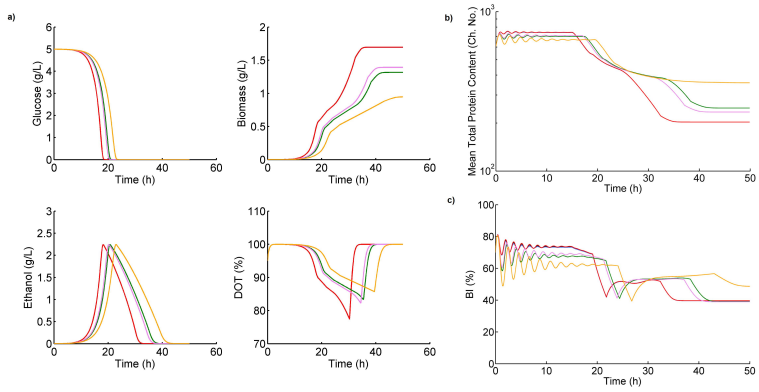

Figure 5.11: Comparison of the model outputs for different combinations of the shape parameters in the partitioning function: $\alpha = \beta = 50$ (black), $\alpha = \beta = 40$ (blue), $\alpha = \beta = 60$ (red), $\alpha = 40$ and $\beta = 50$ (green), $\alpha = 50$ and $\beta = 60$ (pink), and $\alpha = 40$ and $\beta = 60$ (orange). a) variation of the macroscopic variables: glucose, ethanol, biomass and oxygen; b) variation of the mean total protein content; c) variation of the budding index (BI).The black and the blue lines coincide with the red one.

and single-cell flow cytometry data. The experimental single-cell measurements for DNA and total protein contents, as well as the concentrations of substrate and metabolites were interpreted also in the light of the gene expression studies previously reported in the literature.

The standardized procedures developed allow for identifying trends in the single cell properties (critical sizes) along the cultivation. Although, these procedures require the definition of a threshold (e.g. distance from the peak mode) and this decision may introduce a bias, the same threshold decision is made for all the samples, implying that the bias is systematic. On the contrary, manual gating for each sample would introduce a random bias, while the use of the same fixed gates for all samples would neglect the fact that the overall distribution shifts along time thereby also introducing a varying bias.

Population balance models offer a framework to describe the dynamics of distributed properties in a wide range of applications. Although the first discussion on

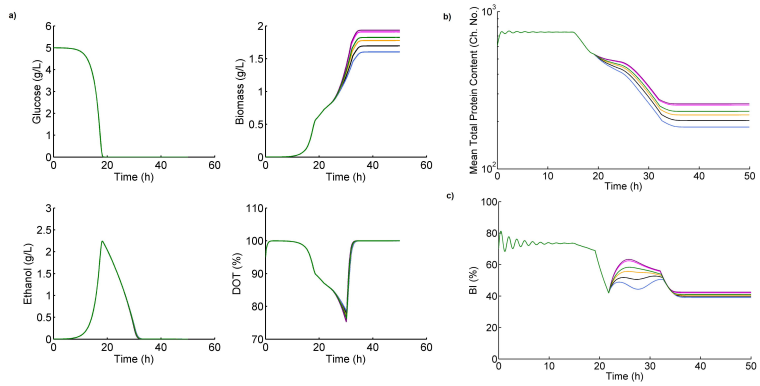

Figure 5.12: Comparison of the model outputs for different combinations of the shape parameters in the partitioning function: $\alpha = 30$ and $\beta = 60$ (black), $\alpha = 20$ and $\beta = 50$ (blue), $\alpha = 30$ $\beta = 50$ (gold), $\alpha = 40$ and $\beta = 50$ (magenta), $\alpha = 50$ and $\beta = 60$ (purple), and $\alpha = 40$ and $\beta = 60$ (green). a) variation of the macroscopic variables: glucose, ethanol, biomass and oxygen; b) variation of the mean total protein content; c) variation of the budding index (BI).The black and the blue lines coincide with the red one.

the use of such models for describing microbial populations was published more than forty years ago [143], reports on the application of the models to specific examples and comparison to experimental results are scarce. One of the causes for such a limited use of these segregated models, in a time where experimental methods for measuring single-cell properties (e.g. flow cytometry, fluorescence microscopy) are easily available, is the difficulty in translating observed behaviors into the PBM growth, budding, division and birth terms.

The overall data analysis allowed for formulating two key kernel functions - budding and division rates - based on dependence of the critical budding and division sizes on the glucose and ethanol uptake rates. Furthermore, the coupling of the PBM to an unstructured kinetic model, proposed in this work, results in a more comprehensive description of the phenomena taking place at different scales (macroscale and microscale) during the cultivation. Such coupling is essential for integrating the typical physiological data (averaged measurements for the overall population), with distributed data collected at single-cell level, thus achieving a model with improved prediction capacities at two levels of detail.

The work here presented contributes towards linking experimental data and PBM theoretical work: a new trend in the PBM community [14]. The proposed model can be regarded as a tool for investigating these dynamics under different scenarios e.g. pulse experiments, and comparing the different assumptions against the experimental observations.

In the case of continuous or fed-batch fermentations where the glucose concentration in the feed is high, the integration of the proposed model and a computational fluid dynamics (CFD) model describing the distribution of substrate within the reactor provides a valuable tool to study in silico the effect of non-ideal mixing, and resulting substrate gradients, on the development of heterogeneous cell populations.

Recently, the occurrence of distribution of protein levels (concentrations) for a cell population was studied *in silico* using a PBM that incorporated a stochastic description for gene expression [197, 198]. It was concluded that, although it is often believed that the occurrence of bimodal distributions results from a bistability of the gene regulatory network (e.g. [199, 200]), this bimodality may arise even if the stochastic bistability does not occur [197]. This interesting observation indicates that bimodality may be a consequence of the dynamics of a population. A PBM including a deterministic, rather than a stochastic description of the protein production kinetics, may eventually be sufficient to describe the distribution of protein levels in a microbial population. Moreover, as the interplay with the extracellular environment (e.g. available substrate) could be included, such a model would be a valuable contribution towards understanding the impact of the development of heterogeneous populations on the overall productivity, and would be a better predictive tool for modeling large-scale fermentors where heterogeneity has been observed [3] and still is a pending problem in practice.

Chapter 6

Population dynamics in a compartmentalized continuously stirred tank reactor

Abstract

In Chapter 5, it was demonstrated that a model consisting of a PBM coupled to an unstructured model was able to capture the development of a budding yeast population, during batch cultivation, in an ideally mixed stirred tank reactor. In this chapter attention is paid to continuous mode cultivations in non-ideally mixed stirred tank reactors. A compartment model approach based on a two compartment continuously stirred reactor is presented and compared to the predicted behavior for a single compartment (ideally mixed) stirred tank reactor. The model framework presented for the batch system Chapter 5 was adapted to a continuously run system, and the critical transition sizes were defined as a function of the available substrate concentration. The effect of reactor compartmentalization on the dynamics of a population of *S. cerevisiae* is assessed, by comparison of cell size and cell cycle position distributions, as well as profiles for macroscopic variables such as glucose, ethanol and oxygen.

6.1 Introduction

As referred in Chapter 2, large scale fed-batch and continuous cultivations are commonly used in industry for the production of biomass and proteins (e.g. insulin). Due to the variation of volume with time and frequent use of varying feed profiles, the interpretation of results from fed-batch cultivations poses additional challenges (in comparison to continuous cultivations). Therefore, continuous cultivations (chemostat) are frequently used in research studies, as well as in industrial scale production.

With regard to continuous cultivations, oscillatory behaviors have been reported for certain combinations of the dilution rate and the dissolved oxygen concentration [201]. Such oscillations are observed for several variables such as substrate (e.g. glucose), metabolites (e.g. ethanol) and budding index (a measure of the distribution of cells over the different cell cycle phases). This type of oscillations is partly related to the cell cycle [201, 202] and the alternation between respiro-fermentative and respiratory metabolisms. Therefore, this phenomenon is commonly designated as respiratory oscillations [202].

In the late 1980s and 90s, many research studies focused on this type of oscillations and different models were proposed (a brief review was made by Bellgardt [122] as well as by Beuse and coworkers [133]). Theoretical work on the induction and mechanisms for sustaining such oscillations were discussed by, for example, Hjortsø and Jensen [119] based on an age-structured PBM and the dependence of the duration of a cell cycle (time to division) on the available substrate.

In particular, a segregated model considering two subpopulations (non-budding and budding cells) was proposed by Cazzador and Mariani [109] (see also [137]). A discretization according to cell mass was applied for each of the subpopulations. The growth (increase in mass) for all cells for a given time interval is assumed to be equal a constant factor, and distinct biomass yields on substrate were assumed for the two subpopulations. Moreover, the extracellular environment is described

by the substrate concentration, and the division of cells (when a critical size is attained) is controlled by a substrate dependent parameter. An empirical mathematical expression is assumed for this dependence. Oscillatory behaviors were successfully reproduced by this model.

In the work presented in this chapter, a PBM coupled to an unstructured model has been used to describe aerobic growth of a population of *S. cerevisiae* in a continuous cultivation. The model proposed in the Chapter 5 is extended in order to describe the supply of substrate and the outflow. In comparison to the model proposed by Cazzador and Mariani [109], the model proposed here includes the following features:

(i) growth of an individual cell depends on cell mass and the extracellular environment;

(ii) ethanol and oxygen are included as model variables, thus extending the description of the extracellular environment;

(iii) two growth modes (on glucose or on ethanol) are considered, and the partitioning of cells (i.e. size ratio between mother and bud) upon division depends on the particular growth mode;

(iv) the critical budding and division sizes are defined as continuous functions of the substrate concentration (glucose or ethanol).

The effect of the operating conditions (dilution rate and glucose feed concentration) on the structure of the population was investigated using the proposed model. In agreement with experimental observations reported in the literature, certain operating conditions lead to oscillating pseudo steady states, while steady state is achieved for other conditions [201]. The oscillations arise from an alternation between growth modes and adjustment of the population in response to decreasing, and then increasing, substrate concentrations. Additionally, the impact of compartmentalization in the bioreactor on the population structure and the occurrence of oscillations is assessed.

Mixing in large scale bioreactors is not ideal, and thus the development of zones, representing different extracellular environments, can occur. For example, the use of Rushton turbines in industrial fermentors (one of the most common impellers used in industrial bioreactors due to the effective gas dispersion) has been observed to generate compartments within the reactor due to the high axial flow barriers created by the turbine [203, 204]. A compartment model approach allows for a crude and simple way of taking spatial heterogeneity in a bioreactor into account, and assessing its impact on the biological phenomena [203]. Additionally, the translation of a compartment model to a laboratory experimental set-up can be easily achieved by using scale-down reactors [4].

In this work, a continuous cultivation of *S. cerevisiae* is described, first for a single compartment model corresponding to the case of an ideally mixed stirred tank reactor. Secondly, a two compartment model was used to describe a continuous cultivation in a spatially heterogeneous stirred tank reactor where a highly concentrated glucose feed is used: one compartment corresponding to the feeding zone and a second one corresponding to the remaining reactor volume. The population dynamics predicted for the homogeneous and for the compartmentalized reactor are compared, and the impact of considering reactor compartmentalization as well as population dynamics (rather than considering an average description of cell behavior, the standard approach in most modeling studies of fermentation processes) to describe the cultivation is discussed.

6.2 Modeling Aspects

6.2.1 Continuously stirred tank reactor (CSTR)

A schematic representation of a continuously stirred tank reactor (CSTR) with a volume V is presented in Figure 6.1 a). The volume is assumed constant, and thus the feed flow rate is equal to the flow rate at the outlet. The dilution rate

($D = F/V$) defines the volume exchange rate. The PBM that was developed for a batch system (Chapter 5) was adapted for the CSTR case (addressed in this chapter) by adding a dilution term in the differential equations corresponding to each cell size pivot and to each of the macroscopic variables described by the unstructured model. For conciseness, the formulation and implementation of the PBM and the unstructured model will be only described for the case of the (more complex) two compartment model. The same principles were used in the case of the single compartment model. In comparison to Chapter 5, a different approach was used for expressing critical budding and division sizes as function of the substrate availability, as further explained in Section 6.2.3.

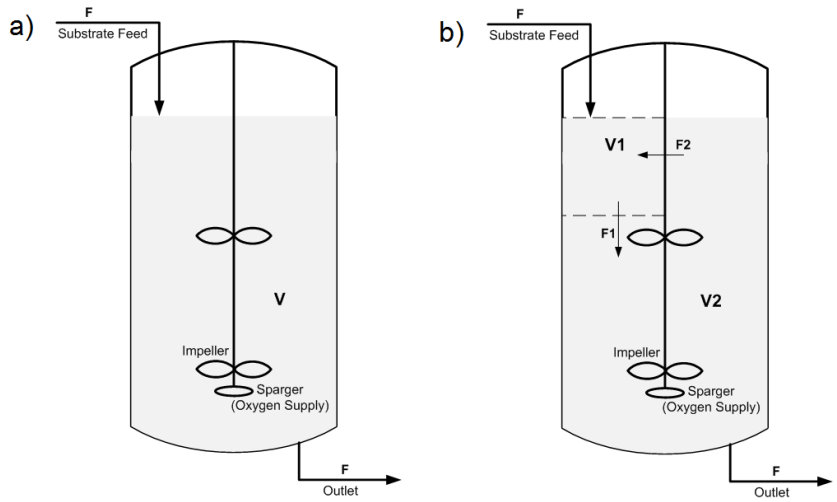

Figure 6.1: Schematic representation of a continuous stirred tank reactor: a) single compartment reactor; b) two compartment reactor.

6.2.2 Two compartment CSTR

In order to represent, in a simplistic fashion, the case of a non-ideally mixed stirred tank reactor, a two compartment reactor was considered. In Figure 6.1 b), a

schematic representation of the compartmentalized bioreactor is provided. A previous study reporting a CFD model for a 22 m^{-3} bioreactor [5], showed clearly that a zone with high substrate concentration was formed in the top of the reactor where the substrate feed was added (between the overhead space and first impeller). Analogously, a smaller compartment, with a volume V1, corresponding to the feeding zone is considered in this study. A larger compartment, with a volume V2, corresponds to the remaining volume of the reactor.

The substrate feed (F), an inflow to compartment V1, consists of a glucose (substrate) rich medium. The internal flows F1 and F2 describe the transport from the compartment V1 to V2, and from compartment V2 to V1, respectively. The outlet of the reactor is located in V2. The oxygen supply is assumed to take place exclusively in compartment V2. Continuous mode operation is assumed, implying that the flow F into V1 is equal to the outlet flow from V2. Assuming constant volumes V1 and V2, the following condition applies: F1= F2+F. For all simulations presented in this chapter, V1 was assumed as 1/6 of the total working volume. As example, a total volume (V) of 30 m^3 was defined (V1 = 5 m^3 and V2 = 25 m^3).

6.2.3 Population Balance Model: compartment model

The population balance equation and boundary conditions for the two stage PBM have been previously discussed in Chapter 5. In this case, four (2 stages x 2 compartments) population balance equations are necessary, and the dilution terms taking into account the transport between compartments and outlet are included in Equations (6.1) to (6.4). The same PBM boundary conditions apply to the two

compartments (Equation (6.5)).

$$\begin{aligned}
\tfrac{\partial N_{V1}^{NB}(m,t)}{\partial t} \quad & + \tfrac{\partial}{\partial m}\left[r_m(m,Z)N_{V1}^{NB}(m,t)\right] = \\
& -\Gamma_B(m\,|Z)\; N_{V1}^{NB}(m,t) \\
& +2\int_{m}^{m_f}\Gamma_D(m'\,|Z)P(m,m'\,\big|Z)N_{V1}^{B}(m',t)dm' \\
& + \tfrac{F2}{V1}N_{V2}^{NB}(m,t) - \tfrac{F1}{V1}N_{V1}^{NB}(m,t)
\end{aligned} \tag{6.1}$$

$$\begin{aligned}
\tfrac{\partial N_{V1}^{B}(m,t)}{\partial t} \quad & + \tfrac{\partial}{\partial m}\left[r_m(m,Z)N_{V1}^{B}(m,t)\right] = \\
& -\Gamma_D(m\,|Z)\; N_{V1}^{B}(m,t) + \Gamma_B(m\,|Z)\; N_{V1}^{NB}(m,t) \\
& + \tfrac{F2}{V1}N_{V2}^{B}(m,t) - \tfrac{F1}{V1}N_{V1}^{B}(m,t)
\end{aligned} \tag{6.2}$$

$$\begin{aligned}
\tfrac{\partial N_{V1}^{NB}(m,t)}{\partial t} \quad & + \tfrac{\partial}{\partial m}\left[r_m(m,Z)N_{V1}^{NB}(m,t)\right] = \\
& -\Gamma_B(m\,|Z)\; N_{V1}^{NB}(m,t) \\
& +2\int_{m}^{m_f}\Gamma_D(m'\,|Z)P(m,m'\,\big|Z)N_{V1}^{B}(m',t)dm' \\
& + \tfrac{F1}{V2}N_{V1}^{NB}(m,t) - \tfrac{F2+F}{V2}N_{V2}^{NB}(m,t)
\end{aligned} \tag{6.3}$$

$$\begin{aligned} \tfrac{\partial N^B_{V2}(m,t)}{\partial t} \quad & + \tfrac{\partial}{\partial m}\left[r_m(m,Z)N^B_{V2}(m,t)\right] = \\ & -\Gamma_D(m\,|Z)\; N^B_{V2}(m,t) + \Gamma_B(m\,|Z)\; N^{NB}_{V2}(m,t) \\ & + \tfrac{F1}{V2}N^B_{V1}(m,t) - \tfrac{F2+F}{V2}N^B_{V2}(m,t) \end{aligned} \tag{6.4}$$

$$\begin{aligned} N^{NB}_{Vi}(m_0,t) = \quad & N^{NB}_{Vi}(m_f,t) = N^B_{Vi}(m_0,t) = {}^B_{Vi}\,(m_f,t) = 0 \\ & m \in [m_0, m_f]; i = 1,2 \end{aligned} \tag{6.5}$$

Critical transition sizes

The critical budding and division sizes (μ_B and μ_D, defining the budding and division rates, Γ_B and Γ_D, respectively) have not been modeled as functions of the substrate uptake rate as previously proposed in the case of the batch cultivation (see Chapter 5). Instead, the critical budding and division sizes were, in this case, considered as continuous functions of the concentrations of glucose or ethanol in the compartment, according to the following assumptions:

(i) If the concentration of glucose, in a given compartment, is equal to or above 0.1 gl^{-1}, glucose growth is assumed for that compartment, and the critical budding (μ_B) and division sizes (μ_D) are calculated based on the glucose concentration, using Equations (6.6) and (6.7). The critical budding and division size maxima are 500 and 950 ch no. and minima of 300 and 550 ch. no., respectively;

$$\mu_B = 200\frac{G}{G + K_{\mu S}} + 300 \tag{6.6}$$

$$\mu_D = 400\frac{G}{G + K_{\mu S}} + 550 \tag{6.7}$$

(ii) If the concentration of glucose, in a given compartment, is below 0.1 gl^{-1}, ethanol growth is assumed for that compartment, and the critical budding (μ_B) and division sizes (μ_D) follow Equations (6.8) and (6.9), corresponding to maxima of 300 and 550 ch no. and minima of 180 and 300 ch. no., respectively;

$$\mu_B = 120\frac{E}{E + K_{\mu S}} + 180 \tag{6.8}$$

$$\mu_D = 250\frac{E}{E + K_{\mu S}} + 300 \tag{6.9}$$

(iii) If the concentrations of glucose and ethanol, in a given compartment, are under 1e-6 gl^{-1}, growth in that compartment is assumed to be equal to zero. An estimated value for the saturation constant of the overall growth process (corresponding to a specific growth rate of half of the maximum value) is 0.15 gl^{-1} [205].

The constants in Equations (6.6) to (6.9) were defined based on the trajectory of the estimated critical sizes along a batch cultivation (see Figure 5.10). The same saturation constant $K_{\mu S}$ was used for all equations and a value of 0.5 gl^{-1} was assumed. The partition shape parameters change according to the glucose and ethanol growth modes as described in Table 6.1.

Table 6.1: PBM partitioning shape parameters for the two growth modes

Partition Parameter	Glucose Growth	Ethanol Growth
Shape parameter α	50	30
Shape parameter β	50	60

Implementation of the PBM

A similar implementation approach as the one described in Chapter 5 was taken. Using the fixed-pivot technique the population balance equations were discretized for each of the 166 pivots. Equations (6.10) and (6.11) are the discretized equations corresponding to Equations (6.1) and (6.2), respectively.

For simplification the terms corresponding to the growth contribution from the neighboring pivots are designated by θ, the negative division and budding term is represented by γ, and the positive budding as birth term by ρ, as marked in Equations (6.10) and (6.11).

In this case, as two compartments and two population states (non-budding and budding) are considered, the ODE system matrix size is 664x644 (166 pivots x 2 compartments x 2 stages = 644 ODEs), as presented in Equation (6.12). The sub-matrices on the top-left and bottom-right corners correspond to the PBM applied to the compartments V1 and V2, respectively, in addition to the outflows for each compartment (see Figure 6.1). The top-right and the bottom-left sub-matrices describe the entrance of the cells due to the recirculation flow from V2 into V1, and from V1 into V2, respectively.

$$\frac{\partial N^{NB}_{V1,m_i}}{\partial t} = \overbrace{\frac{k_m m_i \lambda(Z)}{2(z_i - z_{i-1})} N^{NB}_{V1,m_{i-1}}}^{\theta(m_i)} - \left(\overbrace{\Gamma_B(m_i) + k_m \lambda(Z)}^{\gamma_{NB(m_i)}} + \frac{F1}{V1} \right) N^{NB}_{V1,m_i} - \frac{k_m m_i \lambda(Z)}{2(z_i - z_{i-1})} N^{NB}_{V1,m_{i+1}}$$
$$+2 \sum_{j=i}^{i_{\max}} \left(\underbrace{\left(\int_{m_j}^{m_{j+1}} \frac{m_{j+1} - m}{m_{j+1} - m_j} P(m, m_j) dm + \int_{m_{j-1}}^{m_j} \frac{m - m_{j-1}}{m_j - m_{j-1}} P(m, m_j) dm \right) \Gamma_D(m_j)}_{\rho_{NB}(m_i, m_j)} N^{B}_{V1,m_j} \right) \tag{6.10}$$
$$+ \frac{F2}{V2} N^{NB}_{V2,m_{i+1}}$$

$$\frac{\partial N^{B}_{V1,m_i}}{\partial t} = \overbrace{\frac{k_m m_i \lambda(Z)}{2(z_i - z_{i-1})} N^{B}_{V1,m_{i-1}}}^{\theta(m_i)} - \left(\overbrace{\Gamma_D(m_i) + k_m \lambda(Z)}^{\gamma_{B(m_i)}} + \frac{F1}{V1} \right) N^{B}_{V1,m_i} - \frac{k_m m_i \lambda(Z)}{2(z_i - z_{i-1})} N^{B}_{V1,m_{i+1}} \tag{6.11}$$
$$+ \Gamma_B(m_i) N^{NB}_{V1,m_i} + \frac{F2}{V2} N^{B}_{V2,m_{i+1}}$$

$$
\begin{bmatrix}
\vdots \\
\frac{dN^{NB}_{V1\ m_{i-1}}}{dt} \\
\frac{dN^{NB}_{V1\ m_i}}{dt} \\
\frac{dN^{NB}_{V1\ m_{i+1}}}{dt} \\
\vdots \\
\frac{dN^{B}_{V1m_{i-1}}}{dt} \\
\frac{dN^{B}_{V1m_i}}{dt} \\
\frac{dN^{B}_{V1m_{i+1}}}{dt} \\
\vdots \\
\frac{dN^{NB}_{V2\ m_{i-1}}}{dt} \\
\frac{dN^{NB}_{V2\ m_i}}{dt} \\
\frac{dN^{NB}_{V2\ m_{i+1}}}{dt} \\
\vdots \\
\frac{dN^{B}_{V2\,m_{i-1}}}{dt} \\
\frac{dN^{B}_{V2\,m_i}}{dt} \\
\frac{dN^{B}_{V2\,m_{i+1}}}{dt} \\
\vdots
\end{bmatrix}
=
\left[\begin{array}{ccccccccccc|ccccccccccc}
\ddots & \cdots & 0 & 0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & & \ddots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\theta & -\gamma_{NB}-\frac{F1}{V1} & -\theta & 0 & 0 & 0 & \rho_{NB} & \rho_{NB} & \rho_{NB} & \rho_{NB} & & 0 & F2/V1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \theta & -\gamma_{NB}-\frac{F1}{V1} & -\theta & 0 & 0 & 0 & \rho_{NB} & \rho_{NB} & \rho_{NB} & & 0 & 0 & F2/V1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \theta & -\gamma_{NB}-\frac{F1}{V1} & -\theta & 0 & 0 & 0 & \rho_{NB} & \rho_{NB} & & 0 & 0 & 0 & F2/V1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \cdots & \ddots & 0 & 0 & 0 & 0 & \cdots & & 0 & 0 & 0 & 0 & \ddots & 0 & 0 & 0 & 0 & 0 & 0 \\
\ddots & 0 & 0 & 0 & 0 & \ddots & \cdots & 0 & 0 & 0 & & 0 & 0 & 0 & 0 & 0 & \ddots & 0 & 0 & 0 & 0 & 0 \\
0 & \rho_B & 0 & 0 & 0 & \theta & -\gamma_B-\frac{F1}{V1} & -\theta & 0 & 0 & & 0 & 0 & 0 & 0 & 0 & 0 & F2/V1 & 0 & 0 & 0 & 0 \\
0 & 0 & \rho_B & 0 & 0 & 0 & \theta & -\gamma_B-\frac{F1}{V1} & -\theta & 0 & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & F2/V1 & 0 & 0 & 0 \\
0 & 0 & 0 & \rho_B & 0 & 0 & 0 & \theta & -\gamma_B-\frac{F1}{V1} & -\theta & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & F2/V1 & 0 & 0 \\
0 & 0 & 0 & 0 & \ddots & 0 & 0 & 0 & \cdots & \ddots & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & F2/V1 & \ddots \\
\hline
\ddots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & & \ddots & \cdots & 0 & 0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & F1/V2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & & \theta & -\gamma_{NB}-\frac{F2+F}{V2} & -\theta & 0 & 0 & 0 & \rho_{NB} & \rho_{NB} & \rho_{NB} & \rho_{NB} & \\
0 & 0 & F1/V2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & & 0 & \theta & -\gamma_{NB}-\frac{F2+F}{V2} & -\theta & 0 & 0 & 0 & \rho_{NB} & \rho_{NB} & \rho_{NB} & \\
0 & 0 & 0 & F1/V2 & 0 & 0 & 0 & 0 & 0 & 0 & & 0 & 0 & \theta & -\gamma_{NB}-\frac{F2+F}{V2} & -\theta & 0 & 0 & 0 & \rho_{NB} & \rho_{NB} & \\
0 & 0 & 0 & 0 & \ddots & 0 & 0 & 0 & 0 & 0 & & \cdots & 0 & 0 & \cdots & \ddots & 0 & 0 & 0 & 0 & \cdots & \\
0 & 0 & 0 & 0 & 0 & \ddots & 0 & 0 & 0 & 0 & & \ddots & 0 & 0 & 0 & 0 & \ddots & \cdots & 0 & 0 & 0 & \\
0 & 0 & 0 & 0 & 0 & 0 & F1/V2 & 0 & 0 & 0 & & 0 & \rho_B & 0 & 0 & 0 & \theta & -\gamma_B-\frac{F2+F}{V2} & -\theta & 0 & 0 & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & F1/V2 & 0 & 0 & & 0 & 0 & \rho_B & 0 & 0 & 0 & \theta & -\gamma_B-\frac{F2+F}{V2} & -\theta & 0 & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & F1/V2 & 0 & & 0 & 0 & 0 & \rho_B & 0 & 0 & 0 & \theta & -\gamma_B-\frac{F2+F}{V2} & -\theta & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ddots & & 0 & 0 & 0 & 0 & \ddots & 0 & 0 & 0 & \cdots & \ddots &
\end{array}\right]
\begin{bmatrix}
\vdots \\
N^{NB}_{V1\ m_{i-1}} \\
N^{NB}_{V1\ m_i} \\
N^{NB}_{V1\ m_{i+1}} \\
\vdots \\
N^{B}_{V1m_{i-1}} \\
N^{B}_{V1m_i} \\
N^{B}_{V1m_{i+1}} \\
\vdots \\
N^{NB}_{V2\ m_{i-1}} \\
N^{NB}_{V2\ m_i} \\
N^{NB}_{V2\ m_{i+1}} \\
\vdots \\
N^{B}_{V2\,m_{i-1}} \\
N^{B}_{V2\,m_i} \\
N^{B}_{V2\,m_{i+1}} \\
\vdots
\end{bmatrix}
\tag{6.12}
$$

Coupling to the unstructured model

The unstructured model described in Chapter 5 has been extended to describe the concentrations of glucose, ethanol and oxygen in each of the two compartments (Equations (6.13) to (6.18)). As previously mentioned, the glucose feed is added to compartment V1, and the oxygen transfer is considered to take place exclusively in the compartment V2. As a consequence, the oxygen in V1 is only supplied by the flow F2 and oxygen mass transfer from the gas to the liquid phase *not* is considered in V1.

In a similar fashion to what was described and discussed for the batch system in Chapter 5, the substrate dependent term in the growth kernel, $\lambda(Z)$, is evaluated, in this case, for each of the compartments according to Equation (6.19). Additionally, the biomass concentrations (in grams of dry weight per liter) in each compartment, X_{V1} and X_{V2}, is proportional to the number of cells estimated by the PBM for the respective compartment (see Equation (5.18)).

$$\frac{dG_{V1}}{dt} = -r_{G,\max}\frac{G_{V1}}{G_{V1}+K_G}X_{V1} + \frac{F}{V1}G_{feed} - \frac{F1}{V1}G_{V1} + \frac{F2}{V1}G_{V2} \tag{6.13}$$

$$\frac{dG_{V2}}{dt} = -r_{G,\max}\frac{G_{V2}}{G_{V2}+K_G}X_{V2} - \frac{F+F2}{V2}G_{V2} + \frac{F1}{V2}G_{V1} \tag{6.14}$$

$$\begin{aligned}\frac{dE_{V1}}{dt} = & \overbrace{r_{G,\max}\frac{G_{V1}}{G_{V1}+K_G} - \frac{1}{a}\left(\min\left(r_{O,\max}\frac{O_{V1}}{O_{V1}+K_O}, a\cdot r_{G,\max}\frac{G_{V1}}{G_{V1}+K_G}\right)\right)}^{q_G^{Red}=q_G^{Total}-q_G^{Oxid}}\cdot jX_{V1} \\ & -\frac{1}{k}\left(\min\left(r_{O,\max}\frac{O_{V1}}{O_{V1}+K_O} - \min\left(r_{O,\max}\frac{O_{V1}}{O_{V1}+K_O}, a\cdot r_{G,\max}\frac{G_{V1}}{G_{V1}+K_G}\right), k\cdot r_{E,\max}\frac{E_{V1}}{E_{V1}+K_E}\frac{K_i}{G_{V1}+K_i}\right)\right)X_{V1} \\ & -\frac{F1}{V1}E_{V1} + \frac{F2}{V1}E_{V2}\end{aligned} \tag{6.15}$$

$$\begin{aligned}\frac{dE_{V2}}{dt} = & \overbrace{r_{G,\max}\frac{G_{V2}}{G_{V2}+K_G} - \frac{1}{a}\left(\min\left(r_{O,\max}\frac{O_{V2}}{O_{V2}+K_O}, a\cdot r_{G,\max}\frac{G_{V2}}{G_{V2}+K_G}\right)\right)}^{q_G^{Red}=q_G^{Total}-q_G^{Oxid}}\cdot jX_{V2} \\ & -\frac{1}{k}\left(\min\left(r_{O,\max}\frac{O_{V2}}{O_{V2}+K_O} - \min\left(r_{O,\max}\frac{O_{V2}}{O_{V2}+K_O}, a\cdot r_{G,\max}\frac{G_{V2}}{G_{V2}+K_G}\right), k\cdot r_{E,\max}\frac{E_{V2}}{E_{V2}+K_E}\frac{K_i}{G_{V2}+K_i}\right)\right)X_{V2} \\ & +\frac{F1}{V2}E_{V1} - \frac{F2}{V2}E_{V2}\end{aligned} \tag{6.16}$$

$$\begin{aligned}\frac{dO_{V1}}{dt} = & -\min\left(r_{O,\max}\frac{O_{V1}}{O_{V1}+K_O}, a\cdot r_{G,\max}\frac{G_{V1}}{G_{V1}+K_G}\right)X_{V1} \\ & -\min\left(r_{O,\max}\frac{O_{V1}}{O_{V1}+K_O} - \min\left(r_{O,\max}\frac{O_{V1}}{O_{V1}+K_O}, a\cdot r_{G,\max}\frac{G_{V1}}{G_{V1}+K_G}\right), k\cdot r_{E,\max}\frac{E_{V1}}{E_{V1}+K_E}\frac{K_i}{G_{V1}+K_i}\right)X_{V1} \\ & -\frac{F1}{V1}O_{V1} + \frac{F2}{V1}O_{V2}\end{aligned} \tag{6.17}$$

$$\begin{aligned}\frac{dO_{V2}}{dt} = & k_L a(O^* - O_{V2}) - \min\left(r_{O,\max}\frac{O_{V2}}{O_{V2}+K_O}, a\cdot r_{G,\max}\frac{G_{V2}}{G_{V2}+K_G}\right)X_{V2} \\ & -\min\left(r_{O,\max}\frac{O_{V2}}{O_{V2}+K_O} - \min\left(r_{O,\max}\frac{O_{V2}}{O_{V2}+K_O}, a\cdot r_{G,\max}\frac{G_{V2}}{G_{V2}+K_G}\right), k\cdot r_{E,\max}\frac{E_{V2}}{E_{V2}+K_E}\frac{K_i}{G_{V2}+K_i}\right)X_{V2} \\ & +\frac{F1}{V2}O_{V1} - \frac{F2}{V2}O_{V2}\end{aligned} \tag{6.18}$$

$$\begin{aligned}\lambda(Z)_{V_i} = \quad & \tfrac{b}{a}\left(\min\left(r_{O,\max}\tfrac{O_{V_i}}{O_{V_i}+K_O}, a\cdot r_{G,\max}\tfrac{G_{V_i}}{G_{V_i}+K_G}\right)\right) \\ & +g\left(\overbrace{r_{G,\max}\frac{G_{V_i}}{G_{V_i}+K_G}-\frac{1}{a}\left(\min\left(r_{O,\max}\frac{O_{V_i}}{O_{V_i}+K_O}, a\cdot r_{G,\max}\frac{G_{V_i}}{G_{V_i}+K_G}\right)\right)}^{q_G^{Red}=q_G^{Total}-q_G^{Oxid}}\right) \\ & +\tfrac{l}{k}\left(\min\left(r_{O,\max}\tfrac{O_{V_i}}{O_{V_i}+K_O}-\min\left(r_{O,\max}\tfrac{O_{V_i}}{O_{V_i}+K_O}, a\cdot r_{G,\max}\tfrac{G_{V_i}}{G_{V_i}+K_G}\right), k\cdot r_{E,\max}\tfrac{E_{V_i}}{E_{V_i}+K_E}\tfrac{K_i}{G_{V_i}+K_i}\right)\right), \mathrm{i}=1,2\end{aligned} \tag{6.19}$$

6.2.4 Solution procedure

A solution procedure similar to the one described in Chapter 5 was used. In this case, as two compartments are considered, the growth factor $\lambda(Z)$ is calculated for each compartment (Equation (6.19)) as well as the critical size parameters and partition shape parameters. A schematic representation of the solution procedure used in the work presented in this chapter is provided in Figure 6.2. The MatLab solver ode15s was used for integration of the ODE system (absolute and relative tolerance were set to $1e-3$, while default values were taken for all other solver options).

Due to the dependence of the critical budding and division sizes on the concentration of glucose or ethanol observed at a given time point, the ODE system matrix (Equation (6.12)) has to be evaluated for each iteration of the solution procedure, implying a significant computational effort. For the two compartment model, a simulation for a cultivation of 80h took 5-7 hours computing time using MatLab 2009b release on a 64-bit Windows PC with a 3.30 GHz quad-core processor and 8GB RAM. When the ODE matrix is pre-calculated before entering the iterative loop, the computational time is less than 1 hour (for the same machine).

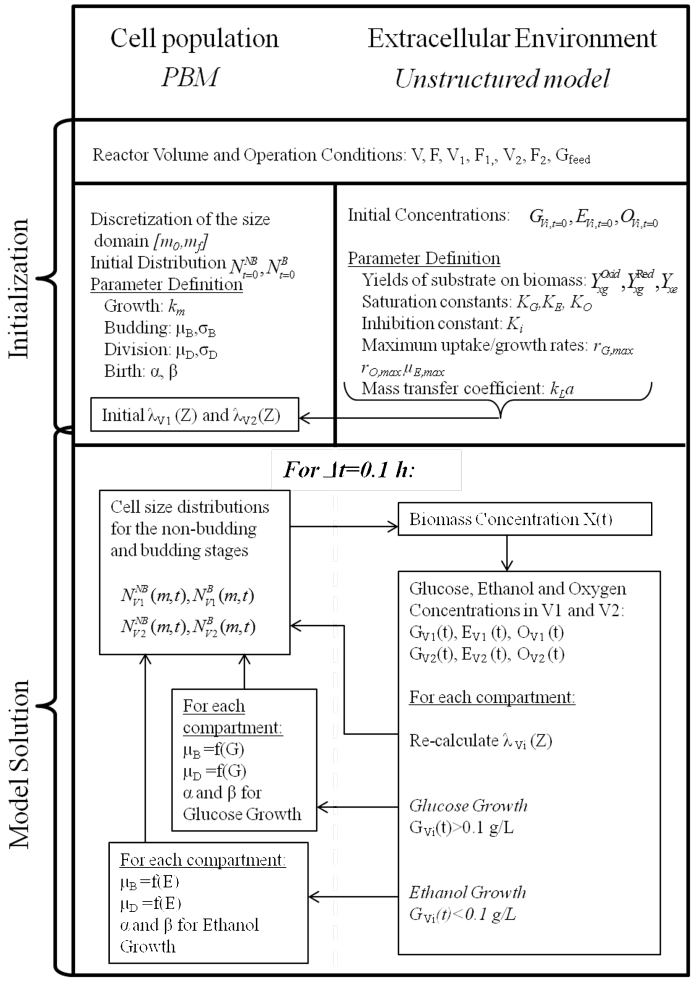

Figure 6.2: Iterative procedured used for solving the two-stage PBM and unstructured model for a two compartment reactor

6.3 Results

The impact of the overall dilution rate ($D = F/V$) and the substrate concentration in the feed flow on the cell size and cell cycle distributions of the budding yeast population was investigated considering both a single compartment reactor and a two compartment reactor. In the latter case, the effect of the recirculation flow (F2) was also investigated. The simulated scenarios correspond to a full 2-level factorial design and the center point. A schematic representation of the 2-factor and 3-factor factorial designs corresponding to the single and two compartment reactor, respectively, is presented in Figure 6.3.

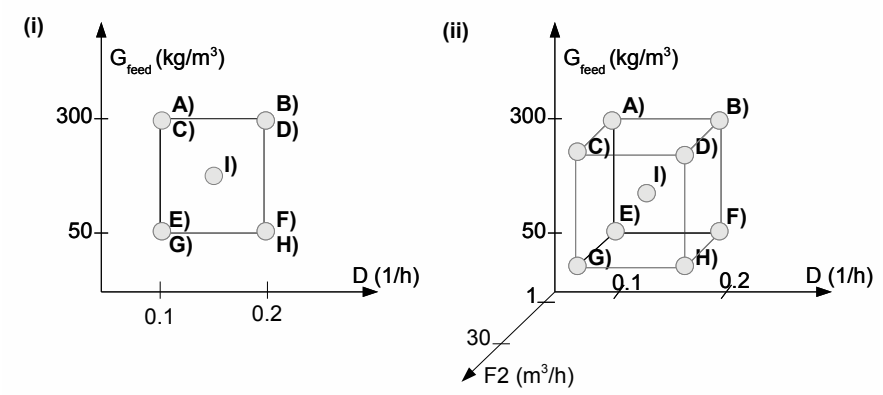

Figure 6.3: Two level factorial design used for evaluating the effect of the overall dilution rate (D) and glucose concentration in the feed flow (G_{feed}), as well as recirculation flow (F2), on yeast population dynamics: (i) 2-factor design for the single compartment reactor), (ii) 3 factor design for the two compartment reactor

6.3.1 Single compartment model

The variation of glucose, ethanol, total biomass (in cell dry weight) along time for the single compartment reactor is presented in Figure 6.4. The cell size dis-

tributions at the simulation end point, for each stage and compartment, are also presented in the same figure.

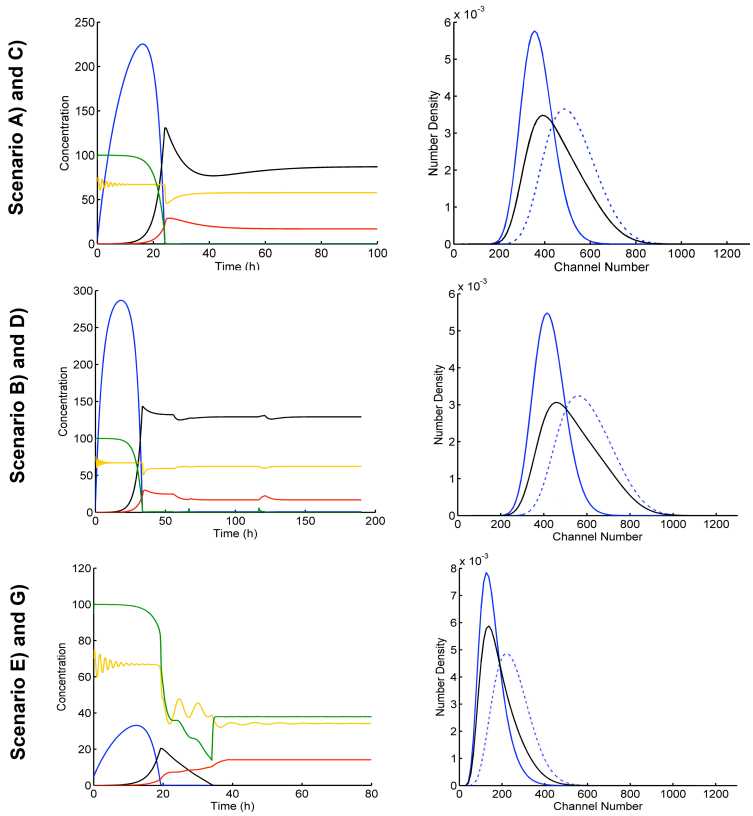

Figure 6.4: Simulation results for scenarios A)-C), B)-D) and E)-G) for the single compartment model: the variation of the concentrations of glucose (blue), ethanol (black), total biomass (red) in gl^{-1}, as well as DO (green) and BI (yellow) in percentage, is presented on the plots on the left-hand side; the normalized cell size distributions for non-budding (full blue line) and budding (dashed blue line) cells, as well as for total cell population (full black line), for the final simulation time, are presented on the plots on the plots right-hand side.

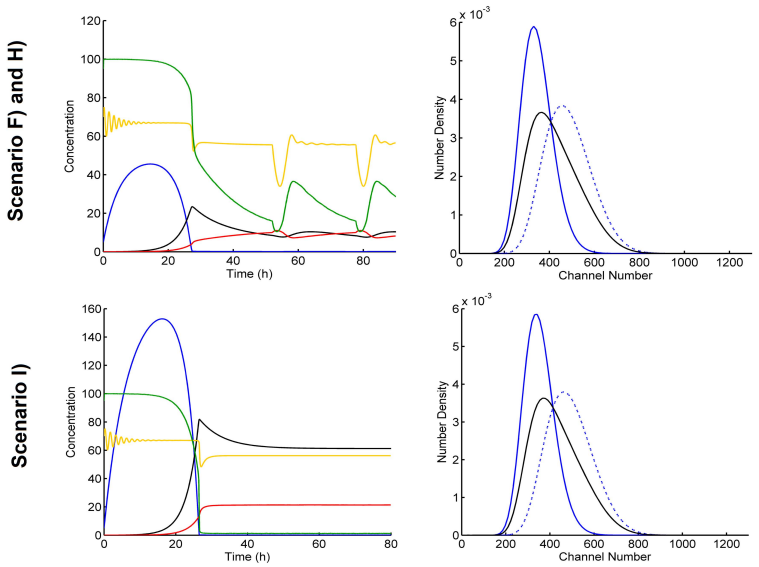

Figure 6.4: (Continued) Simulation results for scenarios F)-H) and I) for the single compartment model: the variation of the concentrations of glucose (blue), ethanol (black), total biomass (red) in gl^{-1}, as well as DO (green) and BI (yellow) in percentage, is presented on the plots on the left-hand side; the normalized cell size distributions for non-budding (full blue line) and budding (dashed blue line) cells, as well as for total cell population (full black line), for the final simulation time, are presented on the plots on the plots right-hand side.

Considering an ideally mixed CSTR (i.e. single compartment), steady state is achieved for all scenarios, except for scenario F)-H) where sustained oscillations are observed. These oscillations are particularly visible in the budding index profiles and oxygen profiles (Figure 6.4). Following an oscillation period, the oxygen is consumed following the consumption of glucose. When the glucose concentration decreases below 0.1 gl^{-1}, the partition coefficients change (switch from glucose growth to ethanol growth) leading to an abrupt increase of the number of newly originated non-budding cells (reflected by the steep fall of the budding index). These small new cells grow slower than larger cells, and thus an accumulation of the glucose and oxygen is observed. When the glucose concentration increases to values above 0.1 gl^{-1}, the partitioning parameters are re-set to the glucose growth mode values, yielding an increase of the budding index.

When comparing various operation conditions, Porro et al [201] suggested that the occurrence of oscillations is observed in a defined range of dilution rates (and thus glucose residual concentrations) and dissolved oxygen concentrations. Such observations are in good agreement with the simulation results here reported.

The ethanol concentration at steady state is only residual for the scenario E)-G). The increased ethanol production observed when the glucose feed concentration increases, and the lower biomass yield when glucose is fermented rather than oxidized, explains the small increase in biomass concentrations observed when comparing scenarios I) and A)-C).

As a result of the residual concentrations of both glucose and ethanol observed at steady state for scenario E)-G), the steady-state cell size distributions are substantially shifted towards smaller cell sizes. Oppositely, the steady state cell size distribution predicted for scenarios B)-D) includes larger cells than observed for other scenarios. This reflects the higher glucose concentration observed at steady state for this scenario.

6.3.2 Two compartment model

When considering a compartmentalized reactor, significant differences, in comparison to the single compartment model, have been observed. The variation of glucose, ethanol, oxygen, total biomass and budding index along time, as well as cell size distributions observed at the simulation end time, are presented in Figure 6.5. Additionally, the concentrations of glucose, ethanol, oxygen and biomass, as well as budding index predicted for all scenarios (at the end simulation time), are compared in Figure 6.6.

Generally, scenarios where a low recirculation flow rate (F2) is imposed (A), B), E) and F)) show the largest differences between compartments in terms of

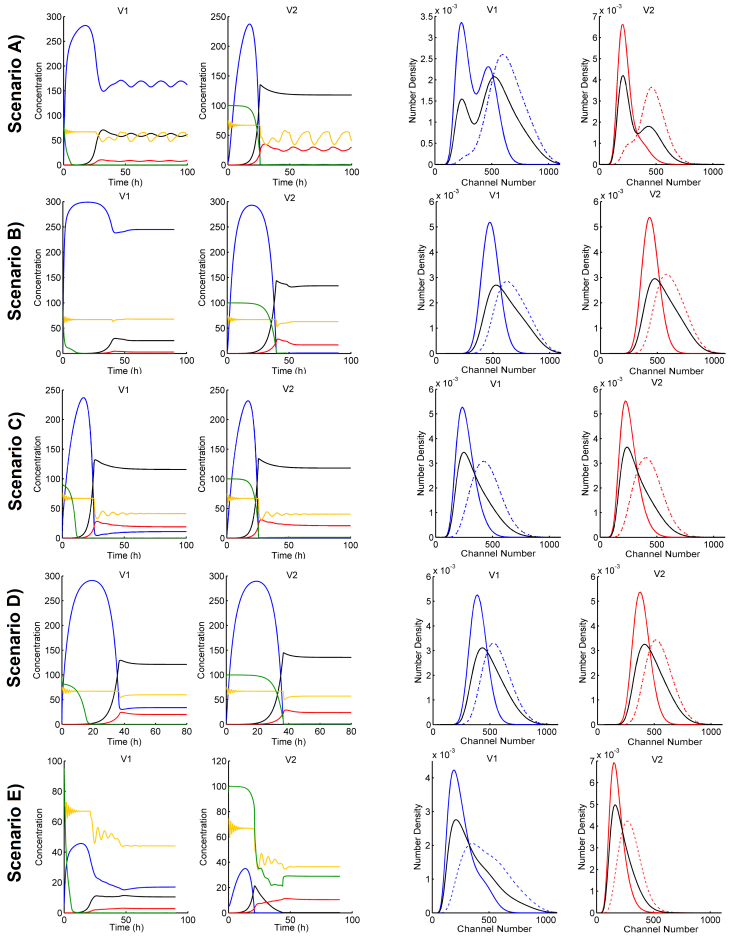

Figure 6.5: Simulation results for scenarios A) to E) for two compartment model: on the left-hand side, the variation of the concentrations of glucose (blue), ethanol (black), total biomass (red) in gl^{-1}, as well as DO (green) and BI (yellow) in percentage; on the right-hand side, the normalized cell size distributions for non-budding (full line) and budding (dashed line) cells, as well as for total cell population (full black line) is presented for the final simulation time.

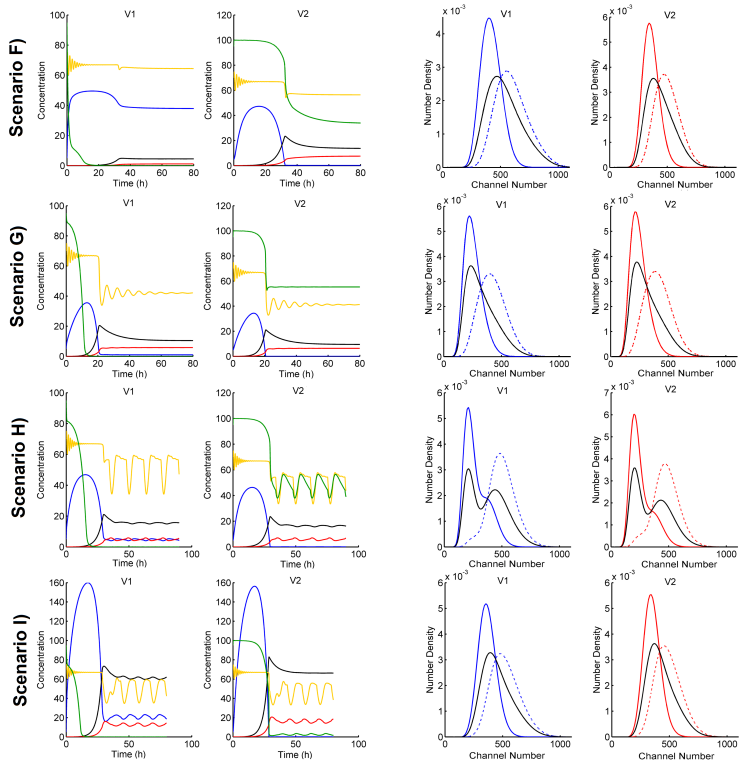

Figure 6.5: (Continued) Simulation results for scenarios F) to I) for two compartment model: on the left-hand side, the variation of the concentrations of glucose (blue), ethanol (black), total biomass (red) in gl^{-1}, as well as DO (green) and BI (yellow) in percentage; on the right-hand side, the normalized cell size distributions for non-budding (full line) and budding (dashed line) cells, as well as for total cell population (full black line) is presented for the final simulation time.

concentrations, but also with regard to predicted budding index and cell size distributions. This is not surprising, as a low recirculation flow rate implies that the liquid exchange between compartments is limited, and, thus, the two compartments show a more independent behavior. Indeed, in the extreme case that F2 would be zero, the two compartment system would, in fact, correspond to two CSTRs operated in series.

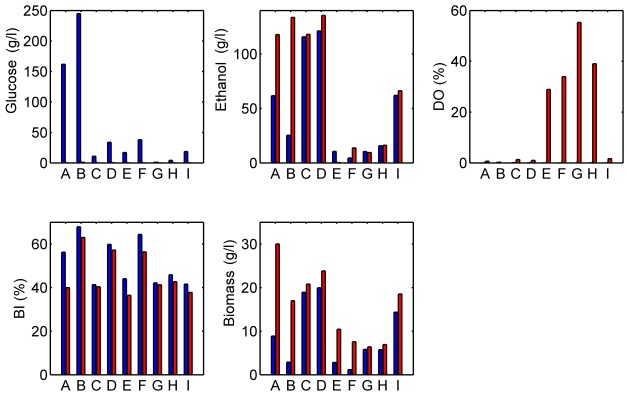

Figure 6.6: Concentrations of glucose, ethanol , dissolved oxygen and biomass, as well as budding index corresponding to scenarios a) to i) for the compartments V1 (blue bars) and V2 (red bars), for the simulation end time.

Residual levels of oxygen in V1, at steady-state, were predicted for all scenarios. Contrarily, the highest dissolved oxygen concentration in compartment V2 is observed for scenario G) (55 % saturation), whereas a residual oxygen concentration was observed for scenario B). The glucose concentration in V1, at steady-state, was higher than threshold value 0.1 gl^{-1} for all scenarios. The lowest concentration was ca. 1 gl^{-1} for scenario G), while the highest concentration was of 245 gl^{-1} for scenario B). The lowest glucose concentration in V2 was also observed in scenario G), and the concentration value (approx. 0.02 gl^{-1}) is under the threshold value of 0.1 gl^{-1}, and thus ethanol growth has been assumed. Glucose concentrations below 0.1 gl^{-1} at steady state are also observed for scenarios C), E). For scenarios A), H) and I), concentrations of glucose, in compartment V2, lower than 0.1 gl^{-1} are instantaneously observed during the oscillation period.

Indeed, oscillatory pseudo-states are observed for scenarios A), H) and I), characterized by simultaneous oscillations of the extracellular environment variables: glucose, ethanol and oxygen. The concentration of ethanol in V2 is an exception in the cases of scenarios A) and I): due perhaps to the high concentrations of

ethanol and very low oxygen availability in compartment V2, it is likely that very small variation in the ethanol concentration takes place and it thus not visible in Figure 6.5.

In case of scenarios A) and I), oscillations were not predicted by the corresponding single compartment model. Oppositely, for scenarios F) and H), oscillations are predicted by both single and two compartment models. This suggests that the existence of sustained oscillations may be scale dependent: oscillatory behaviors observed in laboratory scale experiments may not be transferred to large scale and vice-versa. It seems that the degree of compartmentalization is also a key operational parameter affecting the occurrence of sustained oscillations.While in the case of low dilution rate and high glucose feed concentration (scenario A), a high degree of compartmentalization (low recirculation flow F2) results in oscillatory pseudo-steady state, a lower degree of compartmentalization (high recirculation flow) is necessary for the occurrence of oscillations in the case of higher dilution rate and lower glucose feed concentration (scenario H).

Oscillatory behaviors have been reported in the literature for low dilution rates and/or limited oxygen availability similar to the simulated scenarios A), H) and I) where oscillatory pseudo-steady states have been predicted, not only with regard to oscillation of the glucose concentration but also of the BI and protein content distributions [201].

The variation of the cell size distribution during an oscillation period is illustrated in Figure 6.7. As a result of the continuous dependence of the critical sizes on either the glucose or the ethanol concentration, the adjustment of the cell size distributions is smooth. Similar behavior was observed for scenarios H) and I). When the concentration of glucose in compartment V2 decreases to values under 0.1 gl^{-1}, the ethanol growth mode is considered for this compartment, and the partition shape parameters are set to the values for ethanol growth. In addition, a decrease in the budding and division sizes, that had taken place as a result of the decrease in the glucose concentration (following Equations (6.6) and (6.7)), results in an accumulation of non-budding cells in compartment V2. This is reflected by the increase in the low channel number peak in the non-budding subpopula-

tions (see Figure 6.7, t=50.5 h) and an upwards peak which is visible in the total biomass profile (Figure 6.5).

Due to the transport between compartments, the increase in the number of smaller non-budding cells is also observed for compartment V1. A shift of the cell size distribution of the budding population towards smaller channel no. is then observed, as cells in the non-budding state initiate the budding process (the transition is controlled by a lower critical budding size). Due to the decrease in the growth rate during this period (smaller cells grow slower), and a constant dilution rate, a decrease in the overall biomass concentration, and a corresponding accumulation of glucose, are observed. This accumulation of glucose drives the glucose concentration in V2 above the 0.1 gl^{-1} threshold, and thus the PBM partition parameters are reset to the glucose growth mode value. Parallely, the critical budding and division sizes increase as glucose concentration increases (see Equations (6.6) and (6.7)). Consequently, the population shifts towards large cell sizes, attaining the same distribution observed in the beginning of the oscillation period. The predicted change in the population structure is qualitative agreement with experimental work reported by [201] where the authors observed that the population structure (in terms of protein content distribution) changed during an oscillation period of the dissolved oxygen concentration.

In the case of scenarios D) and F), an adjustment of the budding and division cell sizes in response to the decrease of glucose in both compartments results in a small decrease of the budding index at the time point when the glucose concentration in V2 steeply decreases. Also in the case of scenario G), an adjustment of the budding index as a result of low glucose concentration is observed. However, in this case, the glucose concentration in compartment V2 decreases to values below 0.1 gl^{-1}, implying that partitioning coefficients are set to ethanol growth values, and leading to temporary oscillations that are damped with time.

When comparing scenarios A) and C) (corresponding to the same dilution rate and glucose feed concentration, but different levels for the recirculation flow), a higher biomass concentration is observed at the outlet (i.e. in V2) for the low recirculation scenario: an average biomass concentration of approximately 27 gl^{-1}

(along an oscillation period) is predicted for scenario A), whereas ca. 21 gl^{-1} is predicted for scenario C). A biomass concentration of ca. 17 gl^{-1} is estimated by the corresponding single compartment model. This is rather surprising as it could be expected that higher titers are to be achieved in an ideally mixed reactor (i.e. not compartmentalized). In this case, however, a higher biomass concentration is predicted for the highly compartmentalized scenario A). The oscillatory behavior predicted for this scenario implies that for each oscillation period a number of non-budding cells are generated (as discussed above). As the mass concentration is assumed to be proportional to the number of cells, the predicted biomass concentration increases with the described increase in the number of cells, although these cells belong to smaller size classes. Indeed, when comparing, for scenario A), the distribution of cells predicted for V2 for the two compartment reactor (see Figures 6.5 and 6.7) to the one for the corresponding single compartment model (Figure 6.4), the population is formed by smaller cells in the first case. Moreover, a budding index of approx. 40% is estimated for compartment V2 for scenario A) (as well as C)) whereas, a budding index of 50% is predicted by the single compartment model.

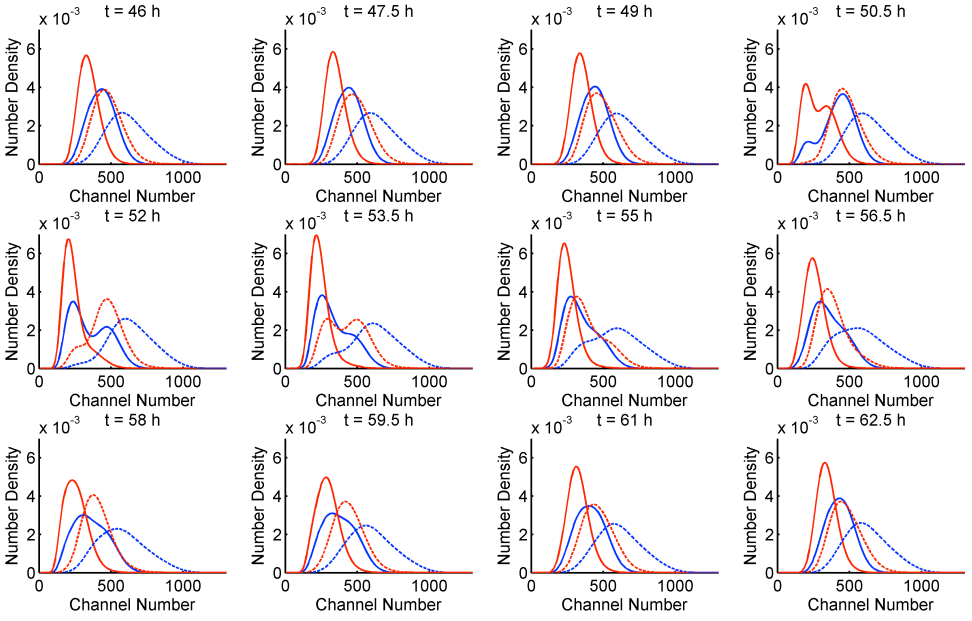

Figure 6.7: Cell size distributions for the non-budding (full lines) and budding (dashed lines) cells populations during an oscillation period, observed in compartments V1 (blue) and V2 (red), for scenario A).

6.3.3 The effect of compartmentalization on overall biomass productivity and yields on substrate

The presented simulation results suggest that the degree of compartmentalization, i.e. spatial heterogeneity, in a bioreactor may have a strong impact on the distribution of cell properties (physiological states) of the cells collected at the outlet of a CSTR. Also significant differences are observed when comparing overall biomass productivity (defined as the product of the biomass concentration at the outlet by the overall dilution rate) and yields of biomass and ethanol (in g per g of consumed glucose) for the one compartment and two compartment models Figure 6.8.

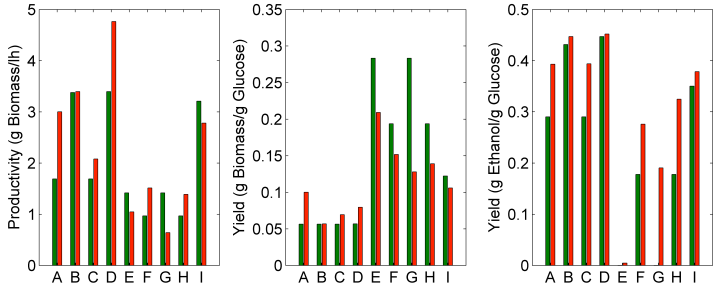

Figure 6.8: Biomass productivity and yield on glucose, as well as the yield of ethanol on consumed glucose for the one compartment (green bars) and two compartment (red bars) models.

For both models, the highest productivity is achieved in scenario D), but due to high ethanol production (and oxygen limitation in the entire reactor) the yield of biomass is relatively low for this scenario. The highest biomass yield is observed for scenario E) where the ethanol concentration (and thus yield of ethanol on glucose) is residual. For the scenarios with lower glucose feed concentration (E), F), G) and H)), the two compartment model predicts lower biomass yields on glucose than the single compartment model, but a higher ethanol yield. In the cases with higher glucose feed concentration (A), B), C) and D)), the two compartment model predicts a slightly higher biomass yield and also a higher ethanol yield. The lat-

ter is particularly visible for scenarios A) and C), characterized by a lower dilution rate. This suggests that the compartment model approach, as oppositely to the single compartment model, takes into account, to larger extent, the occurrence of respiro-fermentative metabolism in certain zones of the reactor yielding a better description of the cultivation.

6.4 Discussion and conclusions

The work presented in this chapter extends the description of a continuous cultivation by predicting the development of cell size distributions for two subpopulations (non-budding and budding), as well as by considering the existence of two compartments within the bioreactor. Cell size and cell cycle position have been used in this work as cell descriptors following the model framework that had been developed for a batch system. These cell properties are easily measured and are deeply connected to growth (as discussed in chapter 5).

The integration of the population balance model and an unstructured model allows for describing the dependence of both the budding and division critical sizes on two different substrates, as well as incorporating the dependence of the partition parameters on the growth conditions. Moreover, the Crabtree effect and the availability of oxygen are accounted for when determining the growth rate.

The proposed extension of the population model framework to a continuous cultivation allowed for reproducing respiratory oscillations that have been reported in the literature for both laboratory and industrial set-ups. Such oscillatory pseudo-steady states could not be reproduced using exclusively a simpler ODE based unstructured model as for example the model proposed by Sonnleitner and Käppeli [182].

A key difference between the formulation for a batch cultivation presented in

Chapter 5 and the formulation for a continuous system presented in this Chapter 6 relies on the mathematical description of the dependence of the critical sizes on the extracellular environment. In the previous chapter, it had been suggested that the adjustment of the critical sizes is triggered based on the overall uptake rate rather than based on a given concentration. The latter was suggested by comparing the data collected experimentally within this thesis work and other results reported in the literature. It is however important to notice that, at this point in time, the complex cell mechanisms underlying the sensing of the substrate availability and consequent adjustment of the regulation of growth (size increase) cell cycle progression (summarized in the adjustment of critical transition sizes) have not been clearly described in a qualitative fashion. Therefore, a quantitative (mathematical) description can only be assumed. In order to explore alternative model formulations, in this chapter the adjustment of the critical sizes was described as a function of the available substrate concentration.

A compartment model approach as the one proposed in this chapter has been proven to be useful for assessing the consequences of reactor compartmentalization due to non-homogeneous mixing patterns. From the study presented in this chapter, it can be concluded that the existence of zones (due to substrate gradients) where cells experience different environments may significantly affect the system behavior in terms of achieving a steady state operation or an oscillatory pseudo steady state, as well as in terms of the predicted overall biomass productivity and yields on glucose. The predicted variations in distributions in the cell size and cell cycle position for various operating conditions, and degrees of compartmentalization, may be perceived as an indicator for other cell properties including the existence of distributions of mRNA levels of specific genes. A previous study [206] on large scale fed-batch cultivation (30 m^3) has shown that, indeed, significant differences in mRNA level of stress genes were observed in different zones in the bioreactor, specially for higher OD levels (i.e. longer cultivation times).

The increase in computational effort resulting from the compartmentalization, although noticeable, is rather acceptable for using the proposed model for testing various scenarios *in silico*. One of the advantages of this type of analysis is that it

is relatively easy to translate the results to a laboratory set-up. In the case of the work presented in this chapter, the two compartment system can be implemented in lab-scale by connecting two stirred tank reactors, as schematically drawn in Figure 6.9. Such a set-up is typically used for scale-down studies in order to simulate large scale conditions in a laboratory set-up (as reviewed by [207]).

Other set-ups, such as a stirred tank reactor connected to a plug-flow reactor, have also been used for scale-down studies (e.g. [4, 206]). However, a major challenge with regard to reproducing large scale conditions is related to the recirculation flows that are imposed between the two bioreactors (compartments). In most reported scale-down studies these flows are assumed without an experimental or theoretical justification. It is, thus, difficult to assess if the flow conditions (and consequently the rate at which cells are being transported between different environments, i.e. compartments) being tested on a scale-down set-up correspond to realistic situations observed in larger scales. Specifically in this regard, the use of computational fluid dynamics (CFD) simulations of flow patterns in large scale reactors may provide a valuable contribution on determining (i) if compartmentalization occurs and for example the relative size of the formed compartments, and (ii) the recirculation flows between existing compartments (e.g. [208]). A review on the use of CFD for characterization and improvement of bioreactor performance can be found elsewhere [160].

In this work, a higher biomass yield was predicted for a highly compartmentalized scenario in comparison to an ideally mixed single reactor. This could suggest that by agitation patterns defining the recirculation flow (and not only agitation speed, as often the description of agitation is reduced to) has a direct impact over the biomass yield. At this point, this suggestion cannot however be regarded as an accurate prediction. In order to be able to develop a model with good predictive power further information on the reactor system would be required. For example, experimental characterization of the scale-down set up in terms of the oxygen transfer coefficient observed in each of the compartment reactors would highly relevant for obtaining a realistic model description of the system.

In conclusion, in this chapter a compartment model approach has been used

to understand the effect of the spatial heterogeneity (i.e. the presence of defined spatial zones corresponding to different extracellular environments) in a bioreactor on the dynamics of a microbial population as well as on the overall system behavior. It is, to my knowledge, the first time such analysis is made based on a population balance model for describing the growth of a microbial population. Analyses as the one here presented and the understanding that can be generated by testing various scenarios *in silico* can be of great value as complementary tool to the experimental scale-down studies.

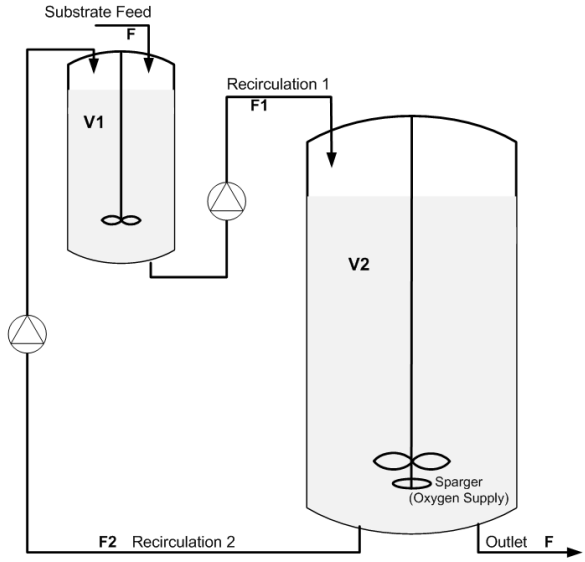

Figure 6.9: Schematic representation of a scale-down reactor set-up consisting of two connected continuously stirred tank reactors.

Chapter 7

Population dynamics in a spatially structured microbioreactor

This chapter consists of an extended version of a peer-reviewed conference contribution:

> Lencastre Fernandes R, Krühne U, Nopens I, Jensen AD, Gernaey KV. *Multi-scale modeling for prediction of distributed cellular properties in response to substrate spatial gradients in a continuously run microreactor*, In: Iftekhar A. Karimi and Rajagopalan Srinivasan, Editor(s), Computer Aided Chemical Engineering, Elsevier, 2012, 31:545-549

Abstract

In large-scale fermentors, non-ideal mixing leads to the development of heterogeneous cell populations [3]. This cell-to-cell variability may explain the differences

in e.g. yields when comparing large- and lab-scale cultivations. In this work, a microscale bioreactor, where substrate gradients occur, is proposed as an *in silico* tool for exploring the complex interplay between flow conditions and population dynamics. This tool provides a proof-of-concept of the integration of computational fluid dynamics (CFD) and population balance models (PBM) for describing microbial cultivations, in spatially heterogeneous bioreactors. The anaerobic growth of *Saccharomyces cerevisiae* in a designed continuously run microbioreactor was modeled. A multi-scale model was developed consisting of a two stage population balance model on the one hand, and a kinetic model integrated into a flow model, on the other hand. The purpose of the model is to predict simultaneously local concentrations of substrate (glucose), product (ethanol) and biomass, as well as the local cell size distributions.

7.1 Introduction

As discussed in Chapter 4, computational fluid dynamics (CFD) is a valuable tool for describing fluid flows in, for example, bioreactors. Such a CFD model allows for a description of the micro-environment that the cells experience at each location in the reactor, and when coupled to kinetic models, the behaviour of cells at each location can be predicted. Moreover, the integration of a PBM into the CFD model would allow for predicting cell property distributions at each location in the reactor. Indeed, by coupling the CFD with the PBM and an unstructured model describing the bulk, a detailed description of the growth conditions and population is obtained for each location in the reactor, as schematically summarized in Figure 7.1.

An Euler-Lagrange approach (see Chapter 4) was applied by Lapin and coworkers [154, 168] for describing lifelines of cells according to their trajectories in a bioreactor. In such an approach, cells are considered as discrete particles (Lagrangian description) that are transported in the continuous phase flow (Eulerian description). A structured model was used to describe the cell metabolism as a function of the extracellular environment, implying that the any cell in the same

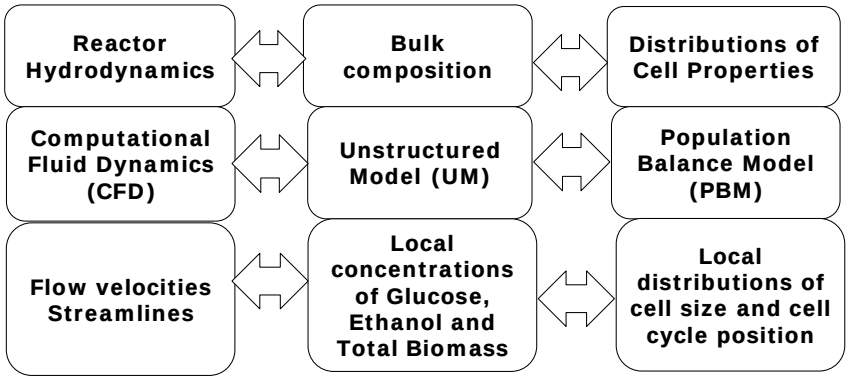

Figure 7.1: Schematic representation of the integration of PBM and CFD models for simultaneous description of the transport throughout the reactor as well as local concentrations and distributions of cell properties.

extracellular environment will behave in a similar fashion.

More recently, taking an Euler-Euler approach, Morchain and co-workers [209, 210] proposed an integration of a PBM, based on the growth rate as cell descriptor variable, and a CFD model. The PBM described the dynamics of a cell population where the overall specific growth rate corresponds to the average growth rate over a distribution of growth rates for individual cells. An increase or decrease of the growth rate of a single cell takes place in response to the specific rate which the extracellular environment allows for: for example, a cell growing at a slower rate than the maximum rate, that could be achieved given the glucose and oxygen concentration in the surrounding medium, will increase its growth rate. In the most recent publication [210], the PBM was integrated into a CFD model for an ideal plug-flow reactor with a residence time of 1.6 h. A good agreement between the profiles predicted by the integrated model and by a corresponding batch model (where only the PBM was considered) was reported.

The main goal of the work presented in this chapter was to develop a simu-

lation tool for *in silico* investigation of the interplay between flow conditions and population dynamics: the result of the transport of substrate and biomass through the reactor, and the dependence of population kinetics on local substrate concentration. Such a tool provides a proof-of-concept of the integration of CFD and PBM that could be used for simulating various flow conditions and kinetic models for the cell population, including cases where spatial heterogeneity (i.e. gradients) is observed.

Despite the significant increase in computational power observed in the last decades, three dimensional (3D) CFD simulations including turbulence models still require simulation times (from weeks to months) that are too long for using such a models as flexible *in silico* simulation tools, where several scenarios and implementations are to be compared in a reasonable time frame. The computational burden is further increased if internal components (monitoring probes, baffles, sparger, etc.) of the bioreactor are included in the CFD geometry and require a refinement of the mesh. Also the integration of kinetic models into the CFD code consists of an additional computation layer that further aggravates the computational time.

In a first effort to address this research problem, it did not seem feasible to start with a complex CFD model for a stirred tank reactor which due to complexity of e.g. rotating parts (the stirrer) would require high computational power and a significant simulation time. Indeed, another aspect that increases significantly the complexity of a CFD model for a bioreactor is the aeration: on the one hand, the presence of a sparger requires the use of a finer mesh (i.e. more detailed and thus having more elements and nodes); on the other hand, describing the presence of a gas phase as bubbles distributed in the liquid phase (i.e. two-phase model) poses considerable challenges and implies a large additional effort from a computational point of view.

Therefore, the focus in this chapter was shifted towards the microscale. Due to the small dimensions, flows in the microscale are laminar and this decreases the complexity of the CFD model, as a turbulence model can be omiited. Moreover, the smaller dimensions imply a smaller number of elements and nodes in the CFD

mesh, thus reducing the computational requirements significantly.

In this work, a continuously run flow-through microbioreactor was designed and the anaerobic growth budding yeast was modeled. By only considering anaerobic growth, the supply of oxygen required for a typically aerobic cultivation is disregarded, further reducing the complexity of the model and , thus, the computational effort. Also the formation and transport of carbon dioxide, as well potential growth inhibition due to its accumulation, were disregarded for simplification purposes.

7.2 Modelling aspects

7.2.1 Flow-through microbioreactor

A microbioreactor was designed for this study. Just as a stirred tank or plug flow reactors, the proposed microreactor is, conceptually, a chamber or reservoir where biomass initially existing in the reactor (by inoculation) grows as a result of the consumption of the substrate provided in the inlet stream. A mixture of biomass, substrate that has not been consumed, and formed metabolites is collected at the outlet. In more detail, the designed microbioreactor consists of a channel (similar to a plug flow reactor) and 5 spherical compartments placed in series around the central channel, as depicted in Figure 7.2. Due to the existence of two symmetry planes (xy and xz planes), only a quarter of the reactor was simulated to further reduce calculation times, and it is represented in the illustrations provided in this chapter.

The spherical compartments were added in order to (i) increase the total volume of the reactor, and (ii) increase the wall surface area implying a decrease of the fluid velocity (as the fluid spreads when entering the spherical compartment).

By increasing the residence time, it is possible to apply higher inlet flow rates (feasible in an experimental set-up) without observing biomass wash out, compared to what would be feasible for a simpler plug-flow reactor without such spherical compartments.

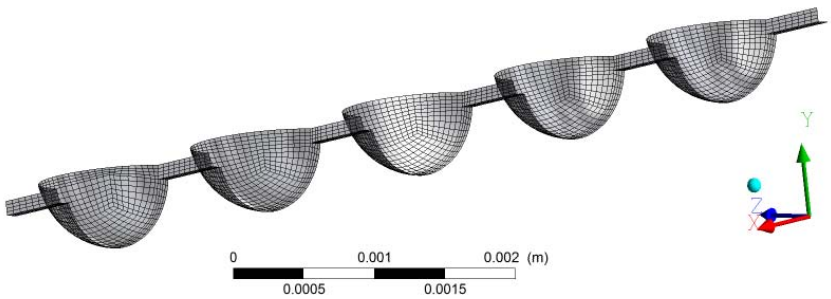

Figure 7.2: Illustration of a section of the designed microbioreactor: the inlet is central channel opening on the left, and the outlet corresponds to the central channel opening on the right. The mesh used for numerical solution of the CFD model is illustrated by the black lines for the reactor wall.

7.2.2 CFD model: Navier-Stokes equation for incompressible fluids

The Navier-Stokes equations describe the velocity or a flow field corresponding to the transport of a fluid in a given volume. Assuming an incompressible fluid, and constant density (ρ) and kinematic viscosity (ν), the Cartesian tensor form of the Navier-Stokes equations is expressed by Equations (7.1) and (7.2), where p stands for pressure and u_i for a component of the velocity vector $u = (u1, u2, u3)$. For the first component, the vector equation is written as Equation (7.3).

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} = -\frac{1}{\rho}\frac{\partial p}{\partial x_i} + \nu \frac{\partial^2 u_i}{\partial x_j \partial x_j} \tag{7.1}$$

$$\frac{\partial u_i}{\partial x_i} = 0 \tag{7.2}$$

$$\frac{\partial u_1}{\partial t} + u_1\frac{\partial u_1}{\partial x_1} + u_2\frac{\partial u_1}{\partial x_2} + u_3\frac{\partial u_1}{\partial x_3} = -\frac{1}{\rho}\frac{\partial p}{\partial x_1} + \nu\left(\frac{\partial^2 u_1}{\partial^2 x_1} + \frac{\partial^2 u_1}{\partial^2 x_2} + \frac{\partial^2 u_1}{\partial^2 x_3}\right) \qquad i = 1 \tag{7.3}$$

At steady state, the transport of a given species A (e.g. glucose or ethanol, cells of a given size) in the flow is described by Equation (7.4), where the left-hand side describes the convective transport, and the right hand side describes the production (sources) or consumption (sinks) of the given species, by for example a (bio)chemical reaction, and the diffusive transport. In this work the diffusive transport has been disregarded for simplification purposes. Numerical solutions of these equations, for the three Cartesian coordinates, using discretization methods as finite-differences, finite-volume or finite elements are normally implemented in the commercial CFD softwares, as the one used in this work.

$$\begin{aligned} u_1\frac{\partial A}{\partial x_1} + u_2\frac{\partial A}{\partial x_2} + u_3\frac{\partial A}{\partial x_3} = \quad & Sources - Sinks \\ & + D_A\left(\frac{\partial^2 A}{\partial^2 x_1} + \frac{\partial^2 A}{\partial^2 x_2} + \frac{\partial^2 A}{\partial^2 x_3}\right) \end{aligned} \tag{7.4}$$

7.2.3 PBM: budding and division critical sizes as functions of the local substrate concentration

The population balance model used in this study consists of the population balance equations and boundary conditions presented in Chapter 5. The partition shape parameters were assumed constant and correspond to the values used in the previous chapters for the growth on glucose: $\alpha = \beta = 50$. In this case, the CFD model accounts for the species transport, and therefore it is not necessary to

include a dilution (transport) term (as it was done in Chapter 6 for a continuously stirred reactor).

The fixed-pivot discretization technique was applied (see Chapter 5) in order to transform the population balance equations into a system of ODEs. In this case, the number of pivots was reduced to 20, in order to decrease the computational effort necessary for the integration with the CFD model (see Section 7.2.5). Hence, the discretized PBM consists of a system of 40 ODEs. Similarly to the work presented in the Chapter 6, the critical budding and division sizes were defined as a function of the local glucose concentrations, following Monod-type kinetics (see Section 6.2). Maximum critical sizes will be considered in glucose abundant locations, whereas the lower critical sizes will be observed for lower glucose concentrations. In locations where glucose is nearly depleted, minimum critical sizes will be used by the PBM.

7.2.4 Unstructured model for the local consumption of glucose and production of ethanol

As described in Chapters 5 and 6 for batch and continuous systems, the PBM was coupled to an unstructured model describing e.g. the consumption of substrate. In this work, the anaerobic growth of the budding yeast (instead of aerobic growth) is considered. The unstructured model, consisting of Equations (7.5) and (7.6), describes the local consumption of glucose, and the local production of ethanol, as a function of the local biomass concentration (estimated based on the total cell number, in a similar fashion to what was described in Chapter 5), and an estimation for the local overall specific growth rate (dependent on the local glucose and ethanol concentrations). The glucose feed at the inlet is accounted for by the fluid

dynamics model, and therefore is not included in the kinetic expressions.

$$\frac{dG}{dt} = -\frac{1}{Y_{XG}}\lambda(Z)X \tag{7.5}$$

$$\frac{dE}{dt} = -\frac{Y_{EG}}{Y_{XG}}\lambda(Z)X \tag{7.6}$$

7.2.5 Implementation of the integrated model in CFX 12.1

The commercial software CFX (v. 12.1, ANSYS Inc., Canonsburg PA, USA) was used in this work. The reactor geometry and a mesh consisting of 32159 hexahedral elements and 36535 nodes (see Figure 7.2), were created using ICEM CFD 12.1, which is part of the CFX software suite.

The ODE system resulting from the discretization of the population balance model, as well as differential equations for the unstructured model were implemented in CFX as user defined expressions using the command expression language (CEL) available in the software.

The CEL language does not offer any advanced mathematical manipulation functions such as numerical integration. Therefore, the elements forming the ODE system (see, for example, Equation (5.1)) for the PBM cannot be evaluated within CFX 12.1. Consequently, the ODE system matrix (e.g. Equation (5.21)) was calculated in MatLab for an array of glucose concentrations ranging from 0 to 100 gl^{-1} (with intervals as small as $1e-3$ for steep regions). For each element in the matrix, a lookup table (saved as a *.tex* file) consisting of the calculated values of

matrix term for the array of glucose concentrations was loaded into CFX. In this way, the CFX model is able to evaluate, for each mesh node, the elements forming the ODEs, according to the concentration of glucose present that given node, by linear interpolation of the look up tables. A schematic description of these steps is provided in Figure 7.3. As mentioned previously in this section, a small number of discretization pivots (20) was used. The reason for selecting such a coarse discretization was the troublesome manual insertion of the equations and time-consuming uploading of the look-up tables into the CFX for each of the elements in the ODE system matrix.

Due to the small dimensions of the designed reactor, a laminar flow regime is expected and, thus, a turbulence model was not used. Transport by diffusion was not considered, and thus all diffusivity coefficients were assumed to be zero.

In order to prevent any fatal errors during the first iterations of the numerical solver, that may occur when solving a complex stiff problem, a steady state solution for the flow problem was obtained before simultaneously solving (numerically) the flow model and the PBM and unstructured models, until convergence of the solution was reached (residuals for the flow model smaller than 10^{-5}, and constant cell size distributions at five different locations in the reactor).

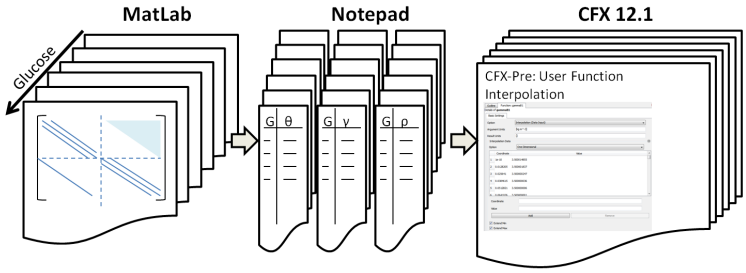

Figure 7.3: Schematic representation of the incorporation of the ODE system matrix terms for various glucose concentrations into CFX Pre 12.1: The matrix terms are calculated for an array of glucose concentrations using MatLab; the resulting array for each matrix term is saved into a text file; for each matrix term a CEL interpolation user function is created in CFX Pre and the corresponding text file is uploaded as a look-up table for interpolation.

Initial conditions

The same initial concentrations of glucose, ethanol and cells of various sizes were assumed for all the locations in the reactor. The initial distributions of non-budding and budding cells correspond to Gaussian distributions with means that approximately correspond to pivots 12 and 14, respectively. The values for cell concentrations with sizes assigned for pivots 1 to 20 are provided in Table 7.1. A budding index of 1% was assumed.

Table 7.1: Initial values for glucose and ethanol concentrations, as well as cell number corresponding to each cell size pivot.

Species	Concentration	Species	Concentration
Glucose	1e-3 gl^{-1}	Ethanol	5e-1 gl^{-1}
Non-budding cells (no. cells l^{-1})		Budding cells (no. cells l^{-1})	
Pivot 1	4.7e5	Pivot 1	7.4
Pivot 2	7.2e6	Pivot 2	3.4e2
Pivot 3	8.1e7	Pivot 3	1.1e4
Pivot 4	6.7e8	Pivot 4	2.8e5
Pivot 5	4.1e9	Pivot 5	5.2e6
Pivot 6	1.9e10	Pivot 6	7.2e7
Pivot 7	6.3e10	Pivot 7	7.2e8
Pivot 8	1.5e11	Pivot 8	5.4e9
Pivot 9	2.8e11	Pivot 9	3.0e10
Pivot 10	8.0e12	Pivot 10	1.2e11
Pivot 11	3.8e13	Pivot 11	3.6e11
Pivot 12	2.9e14	Pivot 12	8.0e11
Pivot 13	1.6e14	Pivot 13	1.3e12
Pivot 14	6.3e13	Pivot 14	1.6e12
Pivot 15	1.9e13	Pivot 15	1.4e12
Pivot 16	4.1e12	Pivot 16	9.4e11
Pivot 17	6.7e11	Pivot 17	4.6e11
Pivot 18	8.0e10	Pivot 18	1.7e11
Pivot 19	7.2e10	Pivot 19	4.5e10
Pivot 20	4.7e9	Pivot 20	8.8e9

Boundary conditions

The outlet was described as an opening, and the software option of *opening pressure and dirn* was selected for boundary details (with a relative pressure of 0 Pa), and the flow direction was set as normal to the boundary condition. The reactor wall was defined as a *no slip wall*.

The inlet flow was defined based on a mass flow rate (see Section 7.3) and the physical properties of water were assumed. A glucose concentration of 20 gl^{-1} was defined. The inlet flow does not contain any cells or ethanol.

7.3 Results

In this work, two different flow conditions were simulated in order to investigate the effect of the flow rate on the cell size and cell cycle distributions at various locations in the microbioreactor. Two inlet flow rates of 1 $ng\ s^{-1}$ and 0.1 $ng\ s^{-1}$ (i.e. approximately 3.6 $\mu l\ h^{-1}$ and 0.36 $\mu l\ h^{-1}$) were selected. Such flow rates are within the operation range of a microfluidic syringe pump (as for example the model NE-1002X commercialized by New Era, Inc. (Farmingdale NY, USA)).

7.3.1 Velocity profiles and streamlines

By solving the Navier-Stokes equations, the three dimensional velocity and pressure profiles can be drawn from the CFD model solution. In Figure 7.4, the flow velocity streamlines[1] are presented for the two flow rate simulations.

[1] Streamlines are a family of curves that are tangent to the velocity vector of the flow. These show the direction a fluid element will travel in at any point in time.

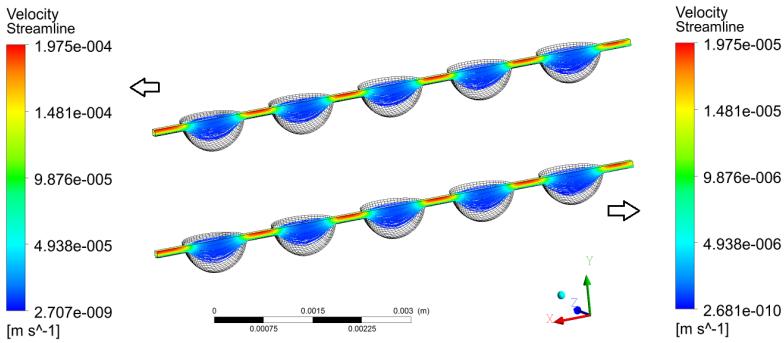

Figure 7.4: Velocity streamlines for the two flow rate conditions: predicted streamlines for the highest flow rate (3.6 $\mu l\ h^{-1}$, top figure) and for lowest flow rate (0.36 $\mu l\ h^{-1}$, bottom figure).

As it could be expected, a difference of an order of magnitude in the flow rates is reflected on an equal difference between the velocities for the two simulation cases. The velocity decreases significantly when the flow enters a spherical compartment and increases in the sections between compartments as pressures increases when the flow is restricted to this central channel. The velocity in the zones close to the compartment walls is nearly zero, and this will promote an accumulation of biomass in these areas, and prevent to some degree a wash out of the biomass.

7.3.2 Glucose, ethanol and total biomass concentration profiles

The glucose concentration profiles for the two simulation cases are presented in Figure 7.5. Glucose is supplied at the inlet of the reactor and is consumed along the reactor as cells use it for growth, resulting in the production of ethanol. Due to the higher residence time for the slowest flow rate (0.36 $\mu l\ h^{-1}$) simulation, the consumption of glucose is higher in this case and the glucose concentration found in the last spherical compartments (closest) to the outlet is residual. Oppositely,

in the case of the higher flow rate, a significant part of the supplied glucose is not consumed, and a high glucose concentration is found at the reactor outlet. In both cases, the glucose concentration at the walls of the spherical compartments is nearly zero, indicating that cells, present in these zones of the reactor, experience glucose depletion.

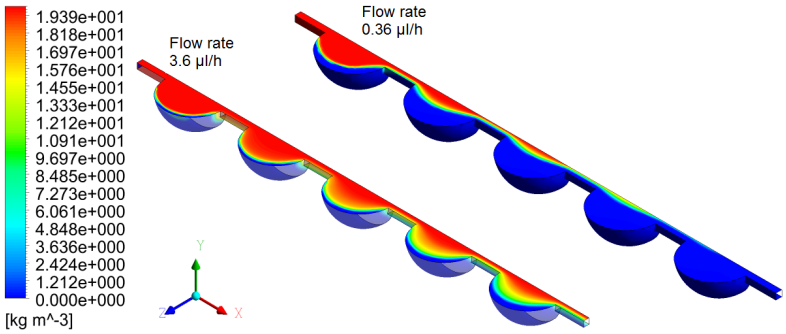

Figure 7.5: Glucose concentration profiles for the two flow rate simulation scenarios. The contour shows the concentrations observed on the reactor wall (spherical compartments and connecting channels walls) and symmetry xz plan.The reactor outlet corresponds to the extremity closest to the bottom right corner.

The predicted profiles for the total biomass concentration are presented in Figure 7.6. Due to the lower flow rate, and thus higher residence time, more biomass is observed for this simulation (0.36 $\mu l\ h^{-1}$) than for the higher flow rate case (3.6 $\mu l\ h^{-1}$). In the latter case, it is clearly visible that biomass does not accumulate evenly on the spherical wall: higher concentrations are found in a wall region closer to the exit of each compartment as illustrated in Figure 7.6. This may be explained by the low flow velocities (nearly zero) observed in these zones, preventing cells from being transported to other parts of the reactor. Cell death is not considered in the model. Therefore, cells in these zones are not growing, and most will remain in these zones unless an alteration in the flow pattern is imposed and they are transported out of these regions. The glucose and ethanol concentrations (Figure 7.7) in these zones are very low, supporting the conclusion that growth does not take place or is residual in these areas.

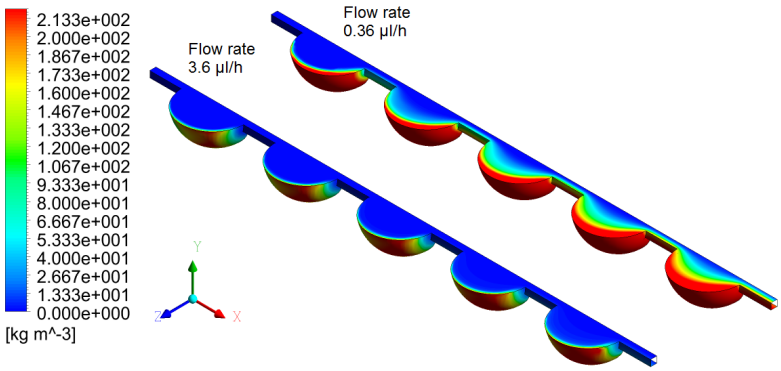

Figure 7.6: Total biomass concentration profiles for the two flow rate simulation scenarios. The contour shows the concentrations observed on the reactor wall (spherical compartments and connecting channels walls) and symmetry xz plan.The reactor outlet corresponds to the extremity closest to the bottom right corner.

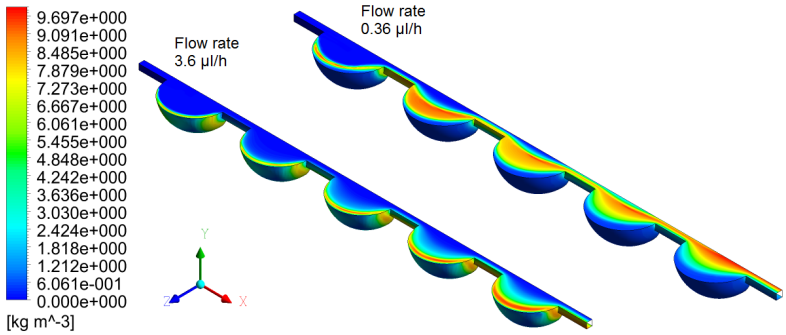

Figure 7.7: Ethanol concentration profiles for the two flow rate simulation scenarios. The contour shows the concentrations observed on the reactor wall (spherical compartments and connecting channels walls) and symmetry xz plan.The reactor outlet corresponds to the extremity closest to the bottom right corner.

Ethanol is formed as glucose is consumed when growth takes place. Not surprisingly, ethanol concentration is highest at the locations where both biomass and glucose are present. When comparing Figures 7.5 to 7.7, it is possible to

observe that generally the highest ethanol concentrations are found in the *interface* between the central stream supplying glucose and the biomass accumulated closer to the spherical compartment walls. However, in the case of the slower flow rate simulation (0.36 $\mu l\ h^{-1}$), a high concentration is observed in the central part of the last spherical compartment, despite the fact that very low growth is expected for this reactor location (low biomass and glucose concentrations). An possible explanation for such high concentration of ethanol in the central streamlines is hydrodynamic focusing: as ethanol formed in other compartments is transported through the reactor, at each pass through a channel connection between spheres, ethanol is concentrated in the central streamlines. Such hydrodynamic focusing may be damped if diffusion had been taken into account in the CFD problem formulation.

7.3.3 Local cell size and cell cycle distributions *vs.* local glucose concentration

In order to investigate locally the interplay of the flow conditions and population dynamics, five locations in the reactor were selected and the cell size distributions for the non-budding and budding cells predicted for each location were compared. The selected locations correspond to five points probes at the same relative position in each of the five spherical compartments as illustrated in Figure 7.8. Due to their similar location relatively to the center of the spherical compartments, the flow velocity, at the five probe points, is the same (1.65 μms^{-1}). The average cell size distributions for the two (cell cycle) subpopulations were, in addition, determined for the reactor outlet plane, and compared to the ones for the probe point locations (Figure 7.9). Further, two additional point probes located on the central channel were defined for the second and the last spherical compartments (B and E, respectively, see Figure 7.8).

For both simulated flow rates, a greater share of smaller cells is found in the compartment closest to the inlet (point A), where the highest glucose concentration is observed. The smaller non-budding cells forming a peak with a mode

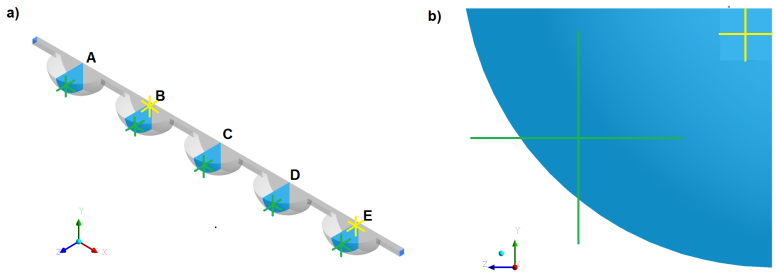

Figure 7.8: Illustration of the selected planes and probe locations: a) five yz planes perpendicular to the central channel were defined at the center of the spherical compartments; b) for each plane a point probe (green) was defined at the same location on the given yz plane. Two additional point probes were defined on planes B and E at a location in the central channel (yellow). The reactor inlet is located at the extremity closest to plane A.

around the pivot 6 (see Figure 7.9 a) and b)), are smaller than the non-budding population given as initial conditions. These cells have most likely resulted from the division of budding cells into non-budding cells with approximately half the size. The bigger cell size peak may correspond to non-budding cells transported from neighboring zones, where, due to lack of glucose, growth is residual and thus cells preserve the characteristics of the initial cell size distribution. The percentage of cells in this bigger cell size peak for the non-budding subpopulation increases for the probe locations closer to the outlet (Figure 7.9 a)), suggesting that a fraction of these non-budding cells in the latter compartments may consist of cells that are transported from previous compartments. In the case of the lower flow rate simulation, the transport contribution to the overall behavior is considerably smaller, and consequently the bigger size peak of non-budding cells (Figure 7.9 b)) is very similar for all locations.

Slightly bigger cells are observed for the budding population (Figure 7.9 c) and d)) in comparison to the non-budding one, as it would be expected based on the model formulation. Similarly to the non-budding cells, two distinct subpopulations (peaks) are also visible for the budding cells population, in both simulation cases, corresponding to a peak with mode around pivot 12, and another peak with mode in pivot 14 (the same mode as the initial distribution of budding cells). In the case

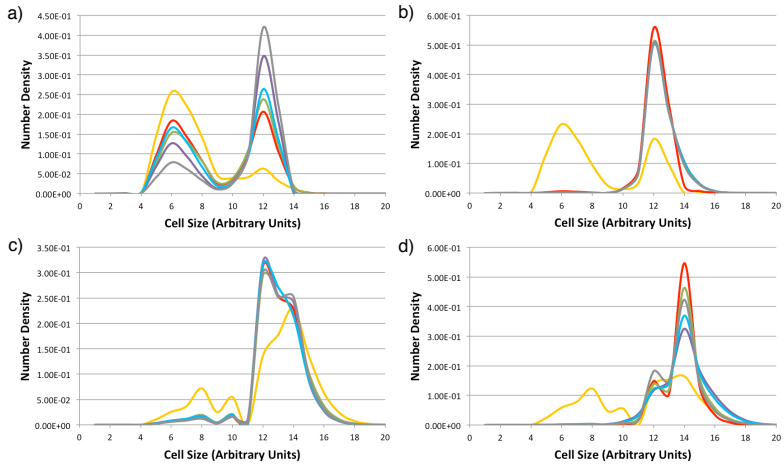

Figure 7.9: Local distributions predicted for the point probe locations closer to the *reactor wall* on plane A (orange), plane B (red), plane C (green), plane D (purple), and plane E (blue), as well as average distribution predicted for the reactor outlet plane (grey): a) Non-budding subpopulation for the high flow (3.6 $\mu l\ h^{-1}$) simulation; b) Non-budding subpopulation for the low flow (0.36 $\mu l\ h^{-1}$) simulation; c) Budding subpopulation for the high flow (3.6 $\mu l\ h^{-1}$) simulation; d) Budding subpopulation for the low flow (0.36 $\mu l\ h^{-1}$) simulation;

of the slower flow rate (0.36 $\mu l\ h^{-1}$), the fraction of budding cells with a bigger size seems to vary with the location (Figure 7.9 d)). As previously mentioned for the non-budding population, these cells are most likely not growing, or growing at a very low rate. A significant decrease in the glucose concentration to near depletion levels is observed, for this slower flow rate, when comparing the first (A) and second (B) compartments. It is, thus, not surprising that the fraction of bigger non-growing budding cells is substantially higher for the second and following compartments.

Considering the two probe locations in plane B (as marked in Figure 7.8), fewer cells are found in the central channel (Table 7.2), and the smaller cell size peak for non-budding cells is not as clearly visible (Figure 7.10 a)) at this location. A similar behavior is observed for the two equivalent locations in plane E (see Figure 7.10

Table 7.2: Concentrations of glucose, ethanol and total biomass (g/l) at the point probe locations, and average concentration predicted for the reactor outlet plane. Points A to E are located closer to the reactor wall on planes A to E. Points B2 and E2 correspond to the point probes located in the central channel on planes B and E, respectively.

	Flow rate: 3.6 $\mu l\ h^{-1}$				Flow rate: 0.36 $\mu l\ h^{-1}$			
	Biomass (gl^{-1})	Glucose (gl^{-1})	Ethanol (gl^{-1})	BI $(\%)$	Biomass (gl^{-1})	Glucose (gl^{-1})	Ethanol (gl^{-1})	BI $(\%)$
Point A	6.7e-4	2.0e1	1.0e-3	70	6.5	1.46	9.0	52
Point B	2.7	1.5e1	2.3	79	1.4	2.9e-25	5.3	19
Point C	6.5	6.7	6.3	78	2.1e2	0	1.6	5
Point D	1.7e1	2.7e-3	9.2	64	2.3e2	0	7.2e-1	2
Point E	1.2e1	1.6e-1	9.2	69	2.3e2	0	8.4e-1	2
Outlet	8.6	9.7	4.8	59	1.52e2	2.8e-8	3.76	6
Point B2	2.6e-2	2.0e1	5.9e-3	33	2.39e-1	1.98e1	1.1e-1	63
Point E2	6.6e1	2.0e-6	2.1e-1	43	1.7e1	1.1e-6	9.1	63

b)). In the case of the lower flow rate simulation (see Figure 7.11), a lower cell size non-budding subpopulation (peak) is however only observed at the probe location closer to the central channel in plane E (closer to the outlet). This implies that, while for the higher flow rate (Figure 7.10 a) and b)) more growing non-budding cells are abundant closer to the wall (smaller size peak), for the slower flow rate these cells can be found in greater share closer to the central channel (Figure 7.11 a) and b)). This can be explained by the larger accumulation of biomass when applying the lower flow rate, implying that the *interface* where cells are present and in contact with glucose rich medium (where ethanol is produced) is located closer to the central channel in the case of the lower flow rate and closer to the wall in the case of the higher flow rate simulation.

Additionally, a larger fraction of smaller budding cells is observed in locations closer to the central channel for both probes in plane B and E in the case of the faster flow rate (Figure 7.10 c) and d)). This indicates that most non-growing (or growing at a very low rate) budding cells are found closer to the wall in both compartments.

In the case of the slower flow rate, there are clearer differences between compartment B and E (Figure 7.10 c) and d)). While in compartment B a smaller fraction of bigger non-growing cells is observed in the central channel in com-

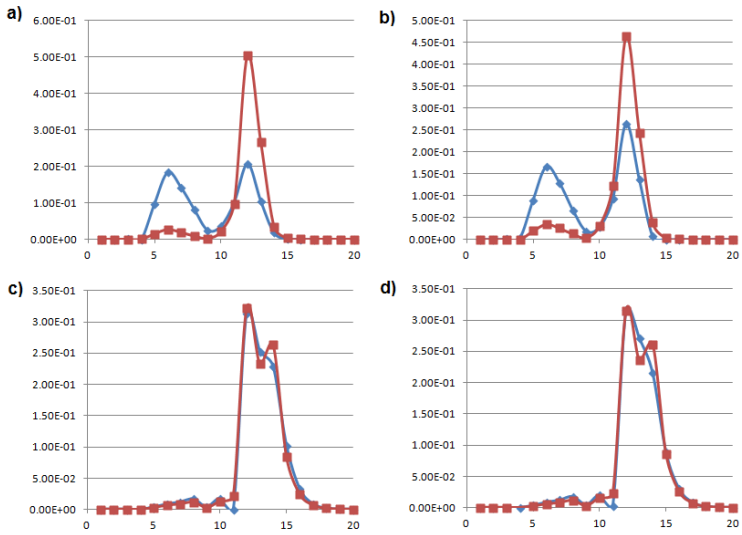

Figure 7.10: Cell size distributions for the higher flow rate simulation (3.6 $\mu l\ h^{-1}$) for: a) the non-budding population at probe location closer to the wall (blue) and in the central channel (red) on plane B; b) the non-budding population at probe location closer to the wall (blue) and in the central channel (red) on plane E; c) the budding population at probe location closer to the wall (blue) and in the central channel (red) on plane B; d) the budding population at probe location closer to the wall (blue) and in the central channel (red) on plane E.

parison to closer to the wall, this difference fades for compartment E. In the latter compartment, a large fraction of non-growing budding cells is observed both closer to central channel and reactor wall. However, a significant fraction of these non-growing cells (in plane E) are bigger cells for the probe location closer to the wall, whereas a significant fraction of smaller budding cells is instead found closer to the central channel (Figure 7.11 d)). This may be explained by the fact that these smaller budding cells are probably originated in upstream compartments, whereas the larger budding cells correspond to the inoculum cells (following the initial distribution) which did not grow.

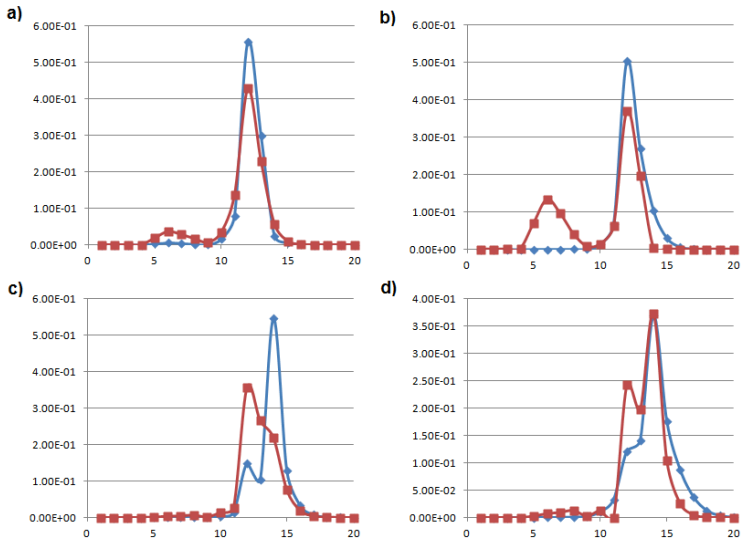

Figure 7.11: Cell size distributions for the lower flow rate simulation (0.36 $\mu l\ h^{-1}$) for: a) the non-budding population at probe location closer to the wall (blue) and in the central channel (red) on plane B; b) the non-budding population at probe location closer to the wall (blue) and in the central channel (red) on plane E; c) the budding population at probe location closer to the wall (blue) and in the central channel (red) on plane B; d) the budding population at probe location closer to the wall (blue) and in the central channel (red) on plane E.

Significant differences in the budding index at the various locations are observed for the two flow rate simulations. A much lower BI is observed for the lower flow rate, especially for locations closer to the wall in the spherical compartments C, D and E (2%, see Table 7.2). Taking into consideration that the BI is only 1% for initial distribution (innoculum), the low BI at steady state, reflects the fact that most cells at these locations where glucose is depleted, do not grow nor undertake the budding and division transitions. Oppositely, in locations close to the central channel, higher budding indexes are observed.

The great difference between the BI values for the central channel probe loca-

tion in Plane E and the outlet can be further understood by analyzing Figure 7.12. As above mentioned, high budding indexes are found in the central fluid stream and decrease steeply when moving towards to reactor wall. At the outlet (similarly to what happens in the channel connections between spherical compartments), due to hydrodynamic focusing, the similar BI contour is observed.

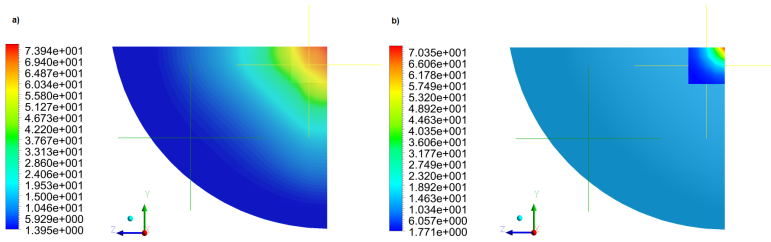

Figure 7.12: Budding Index profiles (in % of budding cells in the total cell population) for a) Plane E and b) the reactor outlet (square section on the top right corner).

In summary, in regions where cell growth takes place (and thus cell cycle transitions occur) relatively high budding indexes are observed and a smaller size subpopulation of non-budding cells is observed. For the lower flow rate, although a significant higher hold-up of biomass within the reactor is achieved. A large fraction of this biomass is not metabolically active (i.e. growing at very low rates, or not growing), due to the existence of spatial glucose gradients. This explains the fact that ethanol concentration at the reactor outlet is in the same order of magnitude for the two flow rate simulations, although the biomass concentration is approximately two orders of magnitude higher for the case of the lower flow rate (0.36 $\mu l\ h^{-1}$, see Table 7.2).

The initial population assumed for the simulation, which would correspond to the inoculum in an experimental cultivation, has a strong impact on the steady state cell size and cell cycle distributions, due to the existence of a non-growing fraction of the overall biomass. In fact, when a BI of approximately 55% is assumed for the innoculum and the same normalized Gaussian initial distribution (similar to the distribution provided in Table 7.1) is used, the budding index observed for the probe locations closer to the wall on planes C, D and E equals

approximately 56%, while in the central channel probe locations on planes B and E, the budding index equals about 50% (data not shown).

7.4 Discussion and Conclusions

The work presented in this chapter provides a proof-of-concept on integration of CFD and a population balance model for describing microbial growth on a non-ideally mixed bioreactor.

Without the coupling of the kinetic models to the CFD model, a comprehensive description of the growth in the reactor, and limitations in the supply of glucose to the accumulated biomass, would not be feasible. A simplified description of the reactor as, for example, a series of plug flow reactors would not be able to account for the fraction of the biomass that is not in contact with media containing glucose and consequently does not grow.

Furthermore, the use of a population balance model provides a detailed insight in the biomass and, as this work demonstrates, subpopulations with various sizes and in different cell cycle phases contribute differently to the total biomass depending on the location in the bioreactor. Consequently, cell growth rate (and metabolic activity) is spatially heterogeneous, and what is observed when collecting samples at the outlet is far from giving a comprehensive understanding of the phenomena taking place inside the bioreactor.

Experimental studies are necessary in order to validate the results here presented and to confirm the usefulness and validity of the results and understanding generated by *in silico* simulations such as the one presented. The construction of the bioreactor proposed in this work by microfabrication methods such as micromilling should not pose substantial problems [211]. Also the collection of samples at the outlet of the reactor for off-line analysis in a flow cytometer should

be straightforward. As described in Chapter 5, there are established methods for measuring cell size and cell cycle distributions for a budding yeast population by flow cytometry. However, the experimental determination of cell size and cell cycle distributions at specific locations in the reactor may be difficult.

Alternatively, a microscope could be used for monitoring. Indeed, due to the small dimensions, the microbioreactor could be used as a typical flow chamber (e.g. [212, 213]) and cultivations could be performed under, for example, a fluorescent microscope where a photography or video camera would be used for monitoring the cultivation. The distribution of glucose within the reactor could perhaps be visualized by using a fluorescent glucose analogue ([214, 215]) as substrate, while the accumulation of biomass (as overall variable and not discriminating subpopulations) could be monitored by use of a reporter system based on a constitutive promoter[2]. More complex reporter systems could be developed in order to monitor specific subpopulations of cells: for example, by targeting promoters of genes specific to the G1 phase of the cell cycle, a fluorescent subpopulation of non-budding cells could be monitored by microscopic analysis.

From a numerical point of view, the use of a discretized grid with only 20 pivots may lead to numerical errors which are difficult to quantify in this work. However, the increase in the number of pivots implies an additional computational effort for a simulation and a cumbersome implementation of the model (pre-processing). When using the CEL language in CFX, equations and user defined functions for parameter interpolation have to be defined manually. Indeed, this is a major drawback. CFX offers a feature for function definition using Fortran routines, but the interface for this feature is far from being straightforward to use. For future exploration of model frameworks as the one proposed, it might be more adequate to find a more flexible programming environment. The use of, for example, COMSOL (COMSOL Inc., Burlington MA, USA), another commercial CFD software, may be more suitable as it provides a link to a advanced mathematical programming software MatLab (MathWorks, Nattick MA, USA). This coupling to MatLab would avoid the use of interpolation tables for each of the system matrix terms. It is often heard in the CFD community that COMSOL still lags behind the ANSYS products,

[2] An unregulated promoter that allows for continual transcription of its associated gene.

CFX or Fluent, for solving complex flow problems. However, in the case of the work here presented, the bottleneck is not the flow problem but the simultaneous solution of a population balance model. Therefore, the integrated use of COMSOL and MatLab may consist of a more adequate alternative to the use of CFX. Another flexible programming environment is offered by Open Foam, an open-source CFD software. In this case however (as it is often for open source software) the degree of user friendliness and the available technical support lag behind the commercial softwares (either COMSOL or ANSYS CFX and Fluent). Further, it should be noted that commercial softwares lag behind on incorporating the newest scientific achievements, and in this regard the use of open source software may be more advantageous.

Despite neglecting diffusive transport of species, which may perhaps have a contribution to the system's behavior that cannot be disregarded, it should be noted that any numerical solution for a fluid dynamics model using an Eulerian formulation (as used in this work) implies a numerical error, due to the truncation error of infinite series used in the discretization of the differential equations. This numerical error is frequently designated as numerical diffusion, as its contribution is similar to the one of diffusive transport [216].

This work has proven that the integration of CFD and population balance models for describing the growth of a microbial population in a spatially heterogeneous reactor is feasible, and that valuable insight on the understanding of the interplay between flow and the dynamics of a budding yeast population (e.g. formation of substrate gradients and non-growth zones) is gained. The development of *in silico* simulation tools may be used for hypothesis generation and testing, and when coupled to an experimental set-up it may be used for process and reactor design optimization.

Chapter 8

Concluding remarks and future perspectives

The work presented in this thesis aimed at proposing modeling approaches in order to describe the dynamics of a heterogeneous microbial population, in response to a varying extracellular environment. More specifically, the project focused on the use of population balance models for describing dynamic distributions of cell properties during cultivation in ideally and non-ideally mixed bioreactors. A compartment model approach and computational fluid dynamics were used for describing non-ideally mixed bioreactors.

Saccharomyces cerevisiae was selected as model organism due to the prior knowledge and know-how from project partners, as well as the vast available literature on *S. cerevisiae* physiology, metabolism and genomics. Cell size and cell cycle were used as descriptors of the yeast population, and thus cell size distributions for non-budding and budding subpopulations were mathematically described by the developed PBM. As discussed in Chapter 4, other cell properties such as age have been used as cell descriptors. The main reasons for selecting cell size and cell cycle were (i) the close link between cell size and cell cycle regulation, in response to variations in the growth conditions, (ii) the existence

of previous literature (although mostly theoretical studies) proposing the use of population balance models based on cell mass (size) as model variable, and (iii) the existence of established methods for measuring these properties in yeast. A strong focus was set, during this project, on mathematically describing the interplay been the regulation single-cell growth and cell cycle progression in response to a changing substrate availability.

Using PBM for describing distributions of cellular properties

The work presented in Chapter 5, takes as starting point a multi-stage PBM using cell size that had been proposed by Hatzis et al [110] for describing the desynchronization of a *S. cerevisiae* population (initially formed exclusively by newly born non-budding cells) under constant growth conditions. In order to establish the critical budding and division sizes as functions of the substrate availability, standardized procedures were here proposed for analysis of the experimentally measured total protein content and DNA content distributions. The observed trends of the critical sizes, along the cultivation, were incorporated in the PBM formulation, and an unstructured model was used to link the dynamics of the distributed properties to the dynamics of extracellular environment variables (glucose, ethanol and oxygen concentrations). The good agreement between the experimental results and the model framework predictions demonstrates that the dynamics of a microbial population in response to decreasing substrate availability can be successfully described in non-stochastic fashion.

Nonetheless, as discussed in Chapter 4, the use of a univariate PBM restricts the description of the metabolic state of individual cells. A multivariate approach using a vector descriptor of the physiological state of the cell could, in theory, provide the most comprehensive description of the population, as well as allow for the identification of sub-populations characterized by co-existing properties (e.g. low concentration of two metabolites and high concentration of a third one). Experimentally, studies using simultaneous monitoring of various mammalian cell properties based on multi-wavelength fluorescent probes have been reported. How-

ever, it must be emphasized that developing methods involving multi-wavelength probes requires among other issues: the accommodation of various lasers in the same equipment, and compensation strategies for dealing to light spill [217].

From a modeling perspective, further challenges arise as the formulation, implementation and solution of a two-dimensional PBM implies a significant additional effort, particularly in the case where the shape of distributions is not known beforehand. In such cases, e.g. of microbial populations, the discretization methods of moments or classes cannot be used without a significant loss of the desired predictive model power. Therefore, a two-stage (non-budding vs. budding) description, instead of considering a second continuous variable, was preferred in the work presented in this thesis. Shu and co-workers [197, 198] have solved a PBM describing the regulated transcription and translation of a protein using a six variable single kinetic model, by transforming the PBM into a stochastic Fokker-Planck equation. Solution of a multivariate deterministic PBM, without projecting the distribution into its moments or requiring other averaging of the distributed cell properties, has not been reported, to the best of my knowledge.

Also the formulation of the multi-dimensional models is not straightforward: it requires a mathematical description of how the various single cell properties depend on each other, as well as of how cell propagation depends on these various properties. Indeed, the formulation of a PBM (one, two or multiple continuous variables) requires an assumption of single-cell kinetics for each variable, and a description of the cell propagation. In the case of the cell size (mass), first-order kinetics on mass, and a multiplicative specific rate dependent on the extracellular environment ($\lambda(Z)$), were assumed for describing the increase of single-cell mass. Cell propagation (division and birth) were defined based on the mass of an individual cell, reflecting the mechanisms of cell size regulation in response to the growth conditions (substrate availability) in the cultivation medium (i.e. existence of critical division sizes) that had been already proposed in the literature.

Assuming single-cell kinetics, for an intracellular cell descriptor (i.e. content of a specific protein, the corresponding mRNA, or the intracellular concentration of a metabolite), and defining mathematically how cells propagate may, however,

not be straightforward. The most obvious choice would be to define single cell kinetics using a structured model describing the average single-cell behavior: for example, using Michaelis-Menten type of kinetics to describe the increase of the single-cell content of a given protein *P* as a function of the intracellular concentration of substrate and decrease due to protein degradation (Equation (8.1), where Y_{PS} is the product yield on substrate, r_{max} is the maximum production rate and k_d is the product degradation rate). The intracellular substrate concentration would depend on the concentration of substrate outside the cell, and the cell mechanisms for transporting it across the cell membrane.

$$\frac{dP}{dt} = Y_{PS} \cdot r_{max} \frac{S}{S + K_S} - k_d \tag{8.1}$$

However, caution should be taken, as the average rates observed for a population may be different from the ones of a single-cell. For example, the same decrease in the average protein content may be explained by a decrease in content in all cells, or the formation of a sub-population with much lower protein content while the remaining cells still present the initial protein content levels (see Figure 1.1). While in the first case a certain average rate might be assumed for all the cells, in the later case a much faster rate for decrease in protein content might be observed for only a fraction of the cells. Indeed, the regulation of certain genes results in switch-like behaviors [218].

Although flow cytometry is a very useful tool for measuring the distribution of cell properties, this technique does not allow for determining kinetics of an individual cell. Indeed, individual cells are not tracked along time when using flow cytometry, but this technique provides time-instant descriptions (snap-shots) of the cell populations. Much progress has been reported with regard to techniques for single-cell measurements; however, the emergence of quantitative measurement methods of single cell dynamics [219] has been very slow. Indeed, the number of quantitative studies yielding single-cell reaction rates is extremely low.

In an interesting study by Zensklusen et al. [220], an *in situ* hybridization approach that detects single mRNA molecules was used for measuring distributions of mRNA copies and transcriptional activity for specific genes in single *S. cerevisiae* cells. Distributions of mRNA copies and description of the expression mode for different genes were evaluated using this technique (the corresponding experimental protocol is provided in another publication [221]). Also many interesting developments in the use of lab-on-chip devices have been reported in the last years (as reviewed in [52]) including time-resolved (which allow to follow a given cell along time) and non-time resolved approaches. Several valuable proof-of-concept studies on applying genomic, proteomic and metabolomic analysis to single cells have been reported. However, as noted by Fritzsch and co-workers [52], such tools remain to be applied to central questions in systems biology. An interesting question to be answered is how the expression of genes is regulated in a single-cell in response to a varying substrate availability. For example, to the best knowledge of the author, there are not any reports on single-cell kinetics for protein production in *S. cerevisiae* under varying growth conditions.

With regard to the mathematical description of the cell propagation, in the cases discussed in this thesis the birth of new cells, by division of a previously existing budding cell, is defined as a function of the PBM variable: the cell size. For other potential cell descriptors (PBM variables), for example the content of a specific protein, cell propagation may not depend on the model variable (i.e. budding cells with any content of the protein of interest undertake division). In this case, a constant propagation rate may be assumed (e.g. [197]). The equal partitioning of protein between the two newborn cells has been assumed in modeling studies for simplicity (e.g. [197]), but any statistical distribution can be assumed (e.g. Gaussian, uniform) to best reproduce experimental observations.

In the Preface of this thesis, it was mentioned that efforts had been made in order to establish a PBM that described the dynamics of the distribution of green fluorescent protein during batch cultivation of a *S. cerevisiae* reporter strain. In this strain, the expression of GFP is controlled by a promoter for the ribosomal protein RPL22a [1]. From the experimental studies, it was possible to conclude that the single expression of GFP decreased significantly during the diauxic shift, and

the content of GFP was proportional to cell size (although the proportionality constant depended on the growth phase). However, it was not possible to evaluate the rate of GFP expression, translation, maturation and degradation, that would be important for describing the single-cell kinetics. Furthermore, frequent sampling and measurement was not feasible, due to logistic problems and the fact that sample freezing resulted in damaged cell membranes and partial leakage of the GFP (and thus the fluorescent signal was altered [1]). Although different theoretical scenarios were simulated, a satisfactory agreement between model predictions and experimentally measured GFP distributions was not achieved.

Similar modeling approaches may be applied for cultivations of other microorganisms or cell lines (e.g. mammalian). A very similar PBM formulation as the one used in this thesis (Chapter 5), based on cell size as model variable, has been used for describing mammalian cell cultivations, and it could, in principle, be applied for any unicellular organism. In the case of, for example *Escherichia coli*, cell cycle progression is also dependent on cells attaining a certain critical size before DNA replication is initiated [222]. Although cells propagate by cell fision (not budding), a two-stage formulation similar to the one used in this thesis (corresponding to different cell cycle phases) could be applied.

For other industrially relevant microorganism such as filamentous fungi, growth and entanglement of the hyphae in pellets, and the number of active hyphae tips, have a significant impact on the performance observed during cultivation. In this case, a PBM where the model variables are the length of the hyphae and the number of actively growing tips may be formulated, as discussed by Nielsen [223]. Liu and co-workers [224] further developed the proposed PBM (based on hyphae length as model variable [223]) by coupling it to a morphological model where different metabolisms (e.g. consumption of glucose, production of streptomycin - a secondary metabolite) are assumed. Parallel to the challenges found during this PhD project, Nielsen [223] identified a need for more fundamental research into the growth kinetics of the individual hyphae and how they interact with each other when forming agglomerates.

Accounting for spatial heterogeneity within a bioreactor

The work described in Chapters 6 and 7 provides an extension of the model framework presented in Chapter 5, for a batch system, to the cases of continuous cultivations in well-mixed and non-ideally mixed bioreactors. Generally, the theoretical studies provided in these two chapters demonstrate that *in silico* simulations may allow for exploring and challenging the understanding of the interplay between flow dynamics, mixing, and population dynamics. Indeed, the interpretation of distributions observed at different locations and assigning causes for the observed phenomena challenges the common understanding which results from analyses performed by scientists trained to analyze and interpret results based on classic average-behavior descriptions of the biological phenomena.

When applying the PBM within a compartment model approach for describing a budding yeast continuous cultivation in a non-ideally mixed stirred tank reactor (Chapter 6), it was possible to observe that the dynamics of the population vary for different operating conditions. Moreover, the interplay between cell size regulation and the substrate availability results in an attractor-repeller system (decrease in substrate availability leads to decrease in size, which then leads to an accumulation of substrate and consequent increase in size) yielding an oscillatory pseudo-steady state for certain operational conditions and degree of compartmentalization. An ODE based model describing the same variables (glucose, ethanol, oxygen and biomass) does not predict oscillatory behaviors for these conditions, for the case of a single or two-compartment model.

Oscillatory behaviors were indeed only predicted for specific combinations of dilution rate, glucose concentration and recirculation flows.This suggests that an optimization of the mixing, often aimed at improving oxygen mass transfer (by exchanging e.g. the type of impellers) may have unexpected consequences on the biological behavior of the system: a variation in the degree of compartmentalization may affect the population structure (and potentially the system performance), despite other controlled operational conditions (dilution rate and feed concentration) being kept constant.

Summarizing, the compartment model has proven to be a simple, but valuable, approach for describing population dynamics in non-ideally mixed stirred tank bioreactors, that does not imply an extraordinary increase in computational effort, as a coupling of PBM to a CFD model would.

In Chapter 7, a proof-of-concept of the integration of a PBM into a CFD model was achieved for describing anaerobic growth of *S. cerevisiae* in a continuously run microbioreactor. Despite the simplifications that were made (a reduced number of cell mass pivots was used for discretization of the PBM, diffusion was disregarded, gas transfer was also neglected), the work presented provides an example of the valuable (although, complex) insight that can be gained when coupling fluid dynamics and biological descriptions of a system. The use of a CFD model for determining flow patterns (velocity fields) enables taking into account phenomena that were disregarded in simpler model approaches. For example, fractions of biomass that are not metabolically active (because the fed substrate does not reach the locations where these cells are found) may be identified, and it may, thus, be possible to explain why the real product yield is lower than the value predicted by other models.

The implementation of CFD models, and further coupling to kinetic descriptions, in larger scales implies an enormous increase in the computational requirements. Typically parallel and cluster computing are used to solve three-dimensional models corresponding to pilot scale and large-scale models. Alternatively graphical processing units (GPU) have been used as replacement for the traditional central processing units (CPU), or in hybrid approaches (e.g. [225]). A decrease of the computational time up to two orders of magnitude has been reported when comparing GPU to CPU processing (e.g. [225, 226, 227]). Nonetheless, such a large increase in performance has not been observed for all cases (e.g. [228]).

The use of such resource- and time-demanding approaches seems only adequate in the cases where the model framework is well defined and has been proven suitable for describing the system. In this regard, various questions concerning the description of the interplay between cell metabolism and extracellular

environment remain to be answered (by experimental hypothesis testing) before the efforts on solving a pilot- or large-scale CFD-PBM model describing microbial growth in a stirred tank reactor seem justifiable. Indeed, such computationally demanding studies will decrease to a minimum the possibility of using the model framework as *in silico* flexible evaluation tool (as the one proposed in Chapter 7).

In the future, it might thus be more suitable, in a first stage, to study separately the flow patterns in large-scale, and use the obtained knowledge to design a *realistic* scale-down set-up. In a second stage, the experimental studies using a smaller scale set-up can be used to generate new understanding of the biological system, that can be translated into modeling tools for scenario evaluation and optimization. For example, a large-scale CFD model for a given stirred tank reactor may be used to extract information on the degree of compartmentalization and recirculation flows between reactors, which can be translated in to laboratory scale by coupling reactors and imposing comparable recirculation flows as determined at larger scales.

Is data analysis and modeling lagging behind experimental developments in single-cell analysis?

It is a reality that, in the last decade, the areas of research and number of studies, where single-cell techniques has been used, have increased significantly. The resulting collection of multivariate data does pose challenges with regard to data visualization and interpretation, and not many new developments have been presented with regard to innovative model approaches. It is, however, my opinion that important biological questions with regard to single-cell behavior during cultivations remain unanswered, and a significant part of the literature focuses solely on reporting single-cell observations and does not go deeper into proposing mechanisms that explain the observed features.

The theoretical modeling basis that exists today for describing distributed cell

properties (with PBM) has been mostly proposed in the 1970s. The link to experimental data and application to realistic problems is, however, still lagging behind. Although the solution of a high-dimensional PBM, providing a more comprehensive description of the population, may still be out of reach in the near future, I believe, based on the work presented, that univariate PBM coupled multivariate descriptions of the extracellular cellular environment, can be successfully be applied for describing microbial population dynamics in response to a varying extracellular environment. The biggest hurdle is still the formulation of single-cell kinetics. This knowledge limitation may only be overcome if quantitative experimental studies are specifically designed to this end, requiring a joint effort between specialists in modeling and experimental work.

Taking cell heterogeneity into account within bioprocess optimization

Despite the renewed interest in the study of cell heterogeneity and increasing number of available experimental techniques for single-cell and population analysis, the knowledge about the key mechanisms in the development of populations (including single-cell kinetics) and the relation between operating conditions and population kinetics, as well as the understanding of the real impact of the degree of cell heterogeneity on the process performance and robustness, are, at this moment, still insufficient in order to be used for process optimization.

From an industrial point of view, although there are several reports on the use of single cell methods such as flow cytometry for monitoring of industrial processes [57], the complexity associated with the analysis of the collected data, and the formulation and solution of segregated models (e.g. PBM), are still, at this moment, large obstacles to the integration of these techniques in industrial common practice. It should be realized as well here that many industrial fermentation production facilities are used for the production of a series of different products, each with a different strain. In such a production plant, efficient use of flow cytometry is even more challenging since data interpretation methods most probably will need

to be adapted to each strain individually.

The overall research projects this PhD was part of (see Preface) have as assumption the existence of an optimal degree of heterogeneity yielding improved process robustness. There is, however, still a significant way ahead until it is possible to define an optimal degree of heterogeneity and corresponding control strategies (in order to run the process at these optimal performance and/or robustness conditions) are proposed. It is, to my opinion, up to the academic research community to continue efforts in order to prove the benefits in applying more sophisticated experimental and modeling population based methods, when addressing industrially relevant challenges. Until then, it is not foreseeable that a *real* paradigm change in the design and optimization of industrial bioprocesses, by monitoring and modeling distributed cell behaviors rather than only averaged ones, will take place.

Appendix A

Included publications

- Lencastre Fernandes R, Nierychlo M, Lundin L, Pedersen AE, Puentes Tellez PE, Dutta A, Carlquist M, Bolic A, Schäpper D, Brunetti AC, Helmark S, Heins A-L, Jensen AD, Nopens I, Rottwitt K, Szita N, van Elsas JD, Nielsen PH, Martinussen J, Sørensen SJ, Lantz AE, Gernaey KV. *Experimental methods and modeling techniques for description of cell population heterogeneity*. (2011) Biotechnology Advances 29:575-599.

- Lencastre Fernandes R, Carlquist M, Lundin L, Heins A.-L., Dutta A , Sørensen SJ, Jensen AD, Nopens I, Eliasson Lantz A, Gernaey KV. *Cell mass and cell cycle dynamics of an asynchronous budding yeast population: experimental observations, flow cytometry data analysis and multi-scale modeling*. Biotechnology and Bioengineering. (2012) (DOI: 10.1002/bit.24749)

- Lencastre Fernandes R, Krühne U, Nopens I, Jensen AD, Gernaey KV. *Multi-scale modeling for prediction of distributed cellular properties in response to substrate spatial gradients in a continuously run microreactor*, In: Iftekhar A. Karimi and Rajagopalan Srinivasan, Editor(s), Computer Aided Chemical Engineering, Elsevier, 2012, 31:545-549

Biotechnology Advances 29 (2011) 575–599

Contents lists available at ScienceDirect

Biotechnology Advances

journal homepage: www.elsevier.com/locate/biotechadv

Research review paper

Experimental methods and modeling techniques for description of cell population heterogeneity

R. Lencastre Fernandes [a], M. Nierychlo [b], L. Lundin [c], A.E. Pedersen [d], P.E. Puentes Tellez [e], A. Dutta [f], M. Carlquist [g], A. Bolic [a], D. Schäpper [a], A.C. Brunetti [h], S. Helmark [i], A.-L. Heins [g], A.D. Jensen [j], I. Nopens [f], K. Rottwitt [h], N. Szita [k], J.D. van Elsas [e], P.H. Nielsen [b], J. Martinussen [d], S.J. Sørensen [c], A.E. Lantz [g], K.V. Gernaey [a,*]

[a] Center for Process Engineering and Technology, Department of Chemical and Biochemical Engineering, Technical University of Denmark, DK 2800 Kgs. Lyngby, Denmark
[b] Department of Biotechnology, Chemistry, and Environmental Engineering, University of Aalborg, Sohngaardsholmsvej 49, DK 9000 Aalborg, Denmark
[c] Molecular Microbial Ecology Group, Department of Biology, University of Copenhagen, Sølvgade 83H, DK 1307K Copenhagen, Denmark
[d] Center for Systems Microbiology, Department of Systems Biology, Technical University of Denmark, DK 2800 Kgs. Lyngby, Denmark
[e] Centre for Ecological and Evolutionary Studies, Department of Microbial Ecology, University of Groningen, Biological Centre, Postbus 14, 9750 AA Haren, Netherlands
[f] BIOMATH, Department of Applied Mathematics, Biometrics and Process Control, Ghent University, Coupure Links 653, B-9000 Ghent, Belgium
[g] Center for Microbial Technology, Department of Systems Biology, Technical University of Denmark, DK 2800 Kgs. Lyngby, Denmark
[h] Department of Photonics Engineering, Technical University of Denmark, DK 2800 Kgs. Lyngby, Denmark
[i] Fermenco ApS, Technical University of Denmark, bldg. 301, DK 2800 Kgs. Lyngby, Denmark
[j] Center for Combustion and Harmful Emission Control, Department of Chemical and Biochemical Engineering, Technical University of Denmark, DK 2800 Kgs. Lyngby, Denmark
[k] The Advanced Centre for Biochemical Engineering, Department of Biochemical Engineering, University College London, Torrington Place, London, WC1E 7JE, United Kingdom

ARTICLE INFO

Article history:
Received 11 October 2010
Received in revised form 4 February 2011
Accepted 31 March 2011
Available online 19 April 2011

Keywords:
Cell heterogeneity
Reporter systems
Flow cytometry
Microscopy
Raman spectroscopy
Microbioreactors
Population balance models
Computational fluid dynamics

ABSTRACT

With the continuous development, in the last decades, of analytical techniques providing complex information at single cell level, the study of cell heterogeneity has been the focus of several research projects within analytical biotechnology. Nonetheless, the complex interplay between environmental changes and cellular responses is yet not fully understood, and the integration of this new knowledge into the strategies for design, operation and control of bioprocesses is far from being an established reality. Indeed, the impact of cell heterogeneity on productivity of large scale cultivations is acknowledged but seldom accounted for. In order to include population heterogeneity mechanisms in the development of novel bioprocess control strategies, a reliable mathematical description of such phenomena has to be developed. With this review, we search to summarize the potential of currently available methods for monitoring cell population heterogeneity as well as model frameworks suitable for describing dynamic heterogeneous cell populations. We will furthermore underline the highly important coordination between experimental and modeling efforts necessary to attain a reliable quantitative description of cell heterogeneity, which is a necessity if such models are to contribute to the development of improved control of bioprocesses.

Contents

* Corresponding author at: Center for Process Engineering and Technology, Department of Chemical and Biochemical Engineering, Building 229, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark. Tel.: +45 45 25 29 70; fax: +45 45 93 29 06.
E-mail address: kvg@kt.dtu.dk (K.V. Gernaey).

0734-9750/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.biotechadv.2011.03.007

1. Introduction

Bioprocessing operations use the activity of living cells for the production of biomass or other products resulting from the cellular metabolism such as proteins, antibiotics or antibodies. In a typical bioreactor operation, cells are kept viable and at a desired state of metabolic activity by adjusting the nutrient levels and reactor variables (e.g. pH and temperature).

As the microbial cultures used in the bioprocess are, most commonly, isogenic, cell heterogeneity would not be expected if the environment was well controlled at adequate conditions. Nonetheless, population heterogeneity has been observed at large scale. In fact, the scale up of bioprocesses from bench-scale to large scale may lead to lower yields and productivities and an increased by-product formation (Bylund et al., 1998; Enfors et al., 2001). This is caused by a decreased capability to maintain a homogeneous environment in large scale bioreactors as compared to well-mixed bench-scale bioreactors used for process development.

Therefore, while the assumption of a perfectly mixed reactor might be realistic for bench-scale reactors, it certainly is not for large scale bioreactors. Due to limited mixing and mass transfer, gradients of, for example, substrate, oxygen and pH are observed in larger reactors (George et al., 1998). In fact, substrate concentrations may range from high concentrations close to the feed port to residual concentrations in zones more distant to this port, the latter caused by different rates of mixing and biological reaction (Larsson et al., 1996). Cells circulating in the reactor are subjected to successively changing conditions, which, by inducing genetic, metabolic and physiological responses, are held responsible for the development of heterogeneous populations. In the past, such populations were shown to present lower productivity than homogeneous ones (Enfors et al., 2001).

The monitoring and control of bioprocesses, found in industry today, does not account for the heterogeneity in microbial populations. The cell properties, determined using on-line, at-line or off-line monitoring methods, correspond to averaged values and, thus, camouflage valuable information on the dynamics of the population (cf. Fig. 1).

Due to the observed decrease in performance at large scale, heterogeneity in bioprocesses was felt to be undesirable. Nonetheless, it might be the key to cell population robustness as observed in tumors (Kitano, 2004) or in cases of bacterial persistence (Balaban et al., 2004). Similar to mechanical stress, which can be exploited to control fungal morphologies to increase overall productivity (Papagianni, 2004), it might also be possible to take advantage of heterogeneity in a microbial population for process optimization. In fact, to understand and harness cell heterogeneity may show us a new path for achieving improved robustness in bioprocesses.

The study of cell heterogeneity in its many aspects has been the focus of the experimental work of many researchers in the recent years, as the number of experimental methods available for single-cell analysis has boomed (Schmid et al., 2010). However, this knowledge has not yet been integrated into a generally accepted modeling framework that is able to account for distributed properties within a cell population, and thus can be used in the design and control of bioprocesses (Müller et al., 2010).

In this contribution, we aim at a) presenting the concept of heterogeneous microbial populations as well as briefly discussing the main factors causing this heterogeneity; b) presenting experimental methods used for studying of cell heterogeneity, in addition to discussing the information within the data sets that these methods yield; c) discussing the design of experiments on microbial cultivations that can deliver information on the development of heterogeneous populations as a result of variations in the extracellular environment;

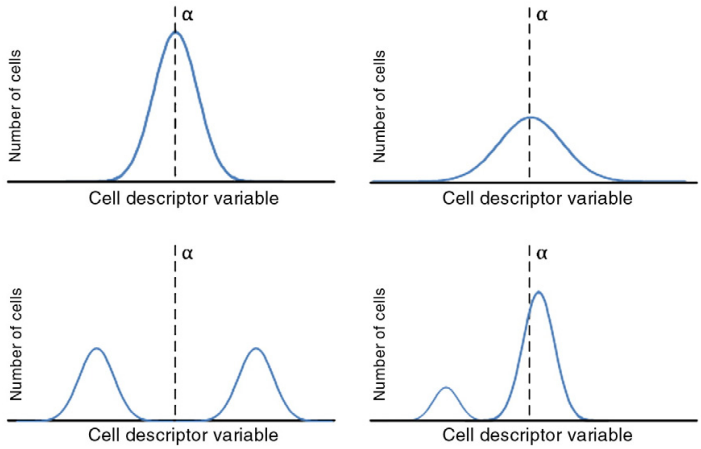

Fig. 1. Description of distributed cell properties. The structure of a population is masked by the use of average cell properties. Different population distributions can correspond to the same mean value α of an experimentally quantifiable cell parameter (e.g. DNA content, NAD(H) concentration, cell mass). (Based on Dhar and McKinney, 2007).

d) presenting and discussing model formulations that can be used for modeling microbial populations, e) and finally, summarizing the potential and future trends in modeling cell heterogeneity as means to increase process knowledge, and, consequently, achieve better bioprocess optimization and control.

2. Cell heterogeneity

Microbial populations have traditionally been thought of as large clonal (i.e. isogenic) groups, thus encompassing identical individuals. However, cell populations in spatially structured environments (like in a large industrial scale bioreactor) can display substantial heterogeneity, where individuals behave differently according to the conditions experienced in their surroundings. This may reflect the different environmental triggers that individual cells experience in a seemingly homogeneous bioreactor (Aertsen and Michiels, 2004). Consequently, the type of environment becomes a key factor for the development of a population and can be considered as a driver of its performance.

2.1. The environment as a driver of microbial population heterogeneity

Generally, environmental stress has been perceived to exert negative effects on microbial populations. In nature, microorganisms are exposed to fluctuating environmental factors, including changes and extremes in temperature, pH, osmolarity, radiation and the concentration of nutrients and toxins. Many microorganisms have developed strategies to cope with such adversities. Central to cell survival are maintaining the integrity of the cell membrane, folding of proteins and the integrity of the DNA (Booth, 2002). To accomplish this, bacterial cells have, for example, developed systems that sense local conditions and determine when these become deleterious and stimulate adaptation. Since response levels may be highest under stress conditions, the control networks have been labeled "stress response" systems (Lengeler et al., 1999; Norman et al., 2005; Ron, 2006). The microbial response to stress is generally accomplished by changes in the expression of those genes whose products are required to combat adversity (Marles-Wright and Lewis, 2007). Some of these stress-induced genes seem to be genuinely specific while others are induced by a wide variety of stresses, and are thus thought to be general stress response genes (De Angelis et al., 2001; Lengeler et al., 1999; Sørensen et al., 2006). Bacteria may use other mechanisms involving physical strategies to survive local adversities, such a sporulation or the use of a flagellum to move to more favorable locations (Serrazanetti et al., 2009). Microbial responses to stress lead to cell differentiation and population fragmentation. In such populations, microbial heterogeneity can be observed.

2.1.1. Biofilms

Biofilms are a clear example of bacterial population differentiation, which, as a result of bacterial specialization yields sub-populations. As the cells adapt to growth in these surface-associated communities, they express phenotypic traits that are distinct from those expressed during planktonic growth (Stewart and Franklin, 2008). Thus, biofilms are spatially structured communities of single or multiple species that show stratified behavior and physiology (Burmølle et al., 2007; Stewart and Franklin, 2008). Within a biofilm, different layered bacterial assemblages form physically and chemically heterogeneous structures with complex architecture. Due to cellular metabolism, this results in the formation of gradients of nutrients, oxygen, waste products, pH, redox potential and electron acceptors. Such factors are spatially influential, resulting in diverse niches on a microscale. This may, at each microsite, select for bacterial variants that are best adapted to the local conditions (Ponciano et al., 2009).

The variability in environmental conditions across space and time (commonly observed in spatially non-homogeneous environments) is called habitat heterogeneity. Thus, environmental conditions and stress are main forces that drive microbial heterogeneity — much like what occurs inside a bioreactor. Biofilm reactors are of extreme relevancy to the understanding of cell heterogeneity, and quite popular in wastewater treatment. We will, however, limit the scope of this review to cultivations of single organisms in suspension. Heterogeneity in biofilms has been reviewed elsewhere (Stewart and Franklin, 2008; Wimpenny et al., 2000) and a general overview on biofilms in wastewater treatment and corresponding models has been edited by Wuertz et al. (2003).

2.2. Responses to environmental stress

The ability of a microorganism to survive stress conditions depends on its genetic content, in which the capability to show adaptive responses is stored. The adaptive responses that occur in a population are classified in three ways: (i) responses at the physiological level, in which cells change behavior (e.g. changes in metabolism due to feedback regulation), (ii) responses at the gene

expression level (global regulation and specific operons (gene regulation)), and (iii) those at the genetic level (genetic changes).

Physiological changes under stress conditions may include byproduct accumulation, changes in substrate uptake capacity, reduction of growth and increased requirements for maintenance energy. For example, under carbon starvation *Escherichia coli* cells become smaller and rounded and accumulate glycogen and polyphosphate (Chung et al., 2006). The DNA is condensed and rapid adjustments in metabolism are made. Moreover, the transport of many macromolecular precursors into the cell is shut down, and ribosome synthesis is blocked. All responses are directed towards keeping energy inside the cell (Chung et al., 2006).

Also, under these conditions, signal transduction systems transmit instructions to the cellular transcription/translation machinery to increase the expression of specific proteins that protect the cell from stress (Moat et al., 2002). A complex network of global regulatory systems with a multitude of molecular components ensures a coordinated and effective response. Such regulatory components include DNA, mRNAs, proteins (e.g. DNA- and RNA-binding proteins) and alternative sigma factors (Wick and Egli, 2004). Under stress conditions, different sigma factors with different promoter specificities are induced, resulting in the expression of specific regulons in response to different stresses. σ^S (RpoS) has been identified as a general stress-responsive sigma factor in Gram-negative bacteria. It helps bacteria to respond to stress like that experienced with the entry into stationary phase. RpoS is the master regulator of general stress responses like that seen upon acid exposure, starvation, high osmolarity and high or low temperature (Delvigne et al., 2009).

Diversification of cell types may also result from DNA sequence changes or rearrangements (Stewart and Franklin, 2008). Genetic rearrangements are movements inside the genome, and encompass several different classes of events: deletions, duplications, inversions and translocations. Also, transposable genetic elements (transposons, insertion/IS elements) can mediate rearrangements (Whoriskey et al., 1987). Genetic rearrangements are thought to occur in a stochastic manner; a single cell in a population that underwent a change may create a mutated population with fitter cells more adapted to the environmental conditions.

Studies on *E. coli* reveal a major contribution of transposons and IS elements to mutagenesis (Arber et al., 1994). In studies of microbial evolution, population takeovers due to stationary-phase mutational events have also been observed in bioreactors by using serial transfer of batch cultures (Lenski and Travisano, 1994). These model systems create culture environments that are spatially structured but essentially constant, leading to a selection of specific phenotypes (Finkel and Kolter, 1999).

3. Experimental methods for characterizing and describing microbial population heterogeneity

As discussed in the previous section, cell heterogeneity resulting from an environmental pressure implies the co-existence of cells at different physiological states. Being able to characterize and predict the physiological state of individual cells in a microbial population is of great importance in a biotechnological fermentation as 1) the physiological state of the individual cell is the only factor that determines the yield of any product, provided that the required nutrients are present in non-limiting amounts, and 2) consequently, the knowledge of the physiological state is a prerequisite for tuning fermentation for optimal performance.

This knowledge has traditionally been acquired indirectly, by measuring a number of parameters like pH, cell density, sugar utilization and product formation. However, as the techniques in the field of molecular biology have improved considerably, the physiological state of cells during the fermentation process has been addressed in much greater detail, primarily by addressing the expression of individual genes, either at the global level by analyzing the transcriptome, or by measuring the expression of genes of particular relevance. Furthermore, the number of studies based on methods able to quantify properties of single cells has increased exponentially in the last years (Schmid et al., 2010).

These single-cell level studies generally aim at understanding the mechanisms lying behind the origin of cell heterogeneity, the cause-effect between observed changes in cells and the micro-environmental conditions in the vicinity of these individual cells, as well as the variations in the environment at a macro-scale. Such methodologies may be of physical, chemical and/or molecular nature and involve a broad range of characteristics, which together give information about the response of populations to environmental cues.

In Section 3, we aim at presenting the experimental methods, which can be used for monitoring the physiological state of individual cells during microbial cultivation. A schematic overview of the scope of this section is presented in Fig. 2.

We start by briefly presenting global methods used to assess genetic heterogeneity in microbial populations (cf. Section 3.1). Global genetic methods allow for an analysis of the development of a genetically heterogeneous cell population. However, these methods are time-consuming and, thus, not suitable for on- or at-line monitoring of bioprocesses.

Reporter systems are typically used in order to achieve an alternative and simpler method for monitoring the cell physiological state. In these systems, a reporter gene is used to monitor the expression state of a certain gene, or the location of the corresponding protein within a cell. In Section 3.2, we provide an introduction and discussion on the use of these systems. In cultivations of strains possessing such reporter systems, the heterogeneity of a population (i.e. distribution of the fluorescent compound) is typically monitored by either microscopy (image analysis) or flow cytometry.

In Sections 3.3 and 3.4, we proceed by presenting these two analytical methods. Both microscopy and flow cytometry underwent substantial advances in the last decades, and are nowadays essential tools for monitoring physiological heterogeneity of microbial populations at single cell level. Indeed, a great number of the applications, of both methods, relies on fluorescence monitoring for measuring cellular parameters, such as the case of reporter systems where the cellular component of interest is fluorescent (e.g. reporter proteins such as green fluorescent protein (GFP)). In addition, these methods allow for monitoring other intrinsic cell properties (e.g. cell size), or structural/functional parameters (e.g. membrane integrity, DNA content), by applying different staining procedures.

Various types of spectroscopic methods have also been applied in monitoring of microbial populations. In fact, in order to achieve a shorter development time, and a higher process understanding, the bioprocess industry has been paying increasing attention to a variety of on- and at-line monitoring methods (Dabros et al., 2008). With regard to single-cell analysis, Raman spectroscopy holds promise as it allows for gathering information about the intracellular composition at single-cell level. We briefly address new developments in the use of Raman spectroscopy in bioprocess monitoring in Section 3.5.

3.1. Global methods for identifying genetic and physiological heterogeneity

Different technologies have been applied to study the genetic and physiological variability found as a result of environmental stresses.

Amplified fragment length polymorphism (AFLP) analysis, pulsed-field gel electrophoresis (PFGE) and repetitive DNA element (rep)-PCR are, among others, frequently used methods for analysis of heterogeneity across microbial genomes. Among these, the rep fingerprinting methods are important; they are based on the use of primers that anneal to repetitive elements in bacterial genomes, which occur in multiple copies. It has been demonstrated that rep-PCR

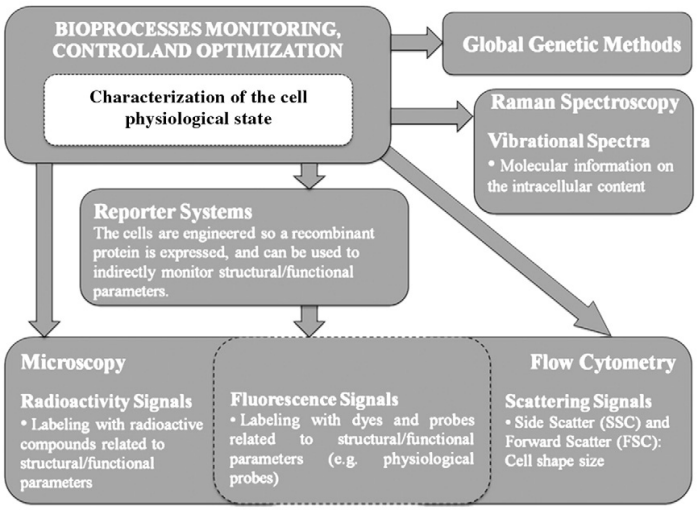

Fig. 2. Monitoring cell physiological states at single-cell level. Methods that can be used for bioprocess control and optimization.

fingerprinting is highly reproducible, being able to distinguish closely related strains, to deduce phylogenetic relationships between strains and to study their diversity in a variety of ecosystems (Rademaker and de Bruijn, 1997). Other techniques, such as suppressive subtractive hybridization (SSH) (which identifies sequences present in one bacterial genome but absent in others) or DNA sequencing are very useful to assess genomic differences between organisms (Beaumont et al., 2009).

Other techniques as DNA microarrays offer a very promising technology to identify microbial heterogeneity. In contrast to conventional studies, which are constrained by a limited number of target genes, microarray-based analysis allows high-throughput analysis and quantification of multiple genes of interest (Wu et al., 2008). Genomic hybridization to a whole-genome microarray detects the presence or absence of similar DNA regions in bacteria, allowing a genome-wide comparison of their genetic contents (Ye et al., 2001). Information about diversity of communities at the level of sub-populations can be obtained when techniques as the ones mentioned above are combined with others such as flow cytometry (e.g. FACS, cf. Section 3.4) which enables subpopulations to be distinguished and physically sorted for further molecular characterization (Lacroix and Yildirim, 2007).

3.2. Reporter systems

The global methods mentioned in the previous section provide the foundation for the selection of specific genes whose expression level confers the relevant information. They are however time-consuming and expensive. As the vast majority of gene products are not easily analyzed, reporter genes are used to confer the required information. Reporter genes typically encode an enzyme whose activity can be assayed easily, either directly (for some reporter systems this can be done on-line) or after cell lysis. The use of reporter genes is one of the oldest tools in molecular biology, initiated with the rIIB gene from the *E. coli* T4 bacteriophage (Champe and Benzer, 1962).

The use of reporter genes is not restricted to confer information of gene transcription frequency (RNA formation), but can also be utilized to provide information about translation, and gene dose (amount of DNA). Moreover, as discussed below, they may also provide information about formation of protein complexes (protein–protein interaction), polypeptide folding, and protein stability. Depending on which parameters are to be addressed, the choice of reporter system should be taken with caution. Reporters are species dependent, as the reporter may act differently in various organisms according to codon usage, translational start signals, protein maturation, and enzymatic parameters.

Several successful applications of reporter systems based for example on the expression of GFP and its variations have been published in the literature. Although traditionally single-cell studies based on reporter systems have been based on microscopic observations, various studies where flow cytometry is used instead have been reported, as reviewed by Davey and Kell (1996), Davey and Winson (2003), as well as Diáz et al. (2010).

3.2.1. Genes encoding β-galactosidase

Since the early 1970s genes encoding β-galactosidases have been extensively used as reporters (Miller et al., 1970; Silhavy and Beckwith, 1985). The first gene was *lacZ* from *E. coli* and this gene is still abundantly used, not only in *E. coli* but also in all cell types like bacteria, yeasts, molds and even higher eukaryotes including human cell lines. Fusions to *lacZ* are typically used to monitor gene expression, either at the level of transcription or translation (Fig. 3). In translational fusions, the promoter and translational initiation signals (start codon and prokaryotic ribosome binding site) of the gene that are going to be monitored are controlling the expression of the reporter, whereas in transcriptional fusions, the translational initiation is controlled by elements that are part of the reporter. These elements need to be optimized for the particular organism. Another issue that needs to be addressed is codon usage. If several rare codons are present in the reporter, they have to be changed according to codon distribution in the host. Neither of the fusions is a measure of the absolute amount of RNA or protein but only relative values as any manipulation of a gene affects its expressions by changing translation efficiency, mRNA half-life, and mRNA maturation and translocation.

Reporters in high-copy-number are not suitable for monitoring gene expression, as they most likely will titrate regulatory proteins resulting in over- or underestimation of the expression level dependent on whether the regulators are activators or repressors. To avoid this, reporters are typically introduced at the chromosome either by homologous recombination at the natural position of the gene of interest or by insertion based on phage or transposon delivery.

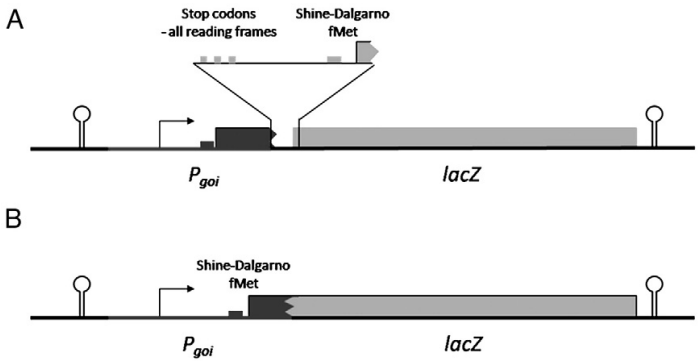

Fig. 3. Genetic maps of transcriptional (A) and translational (B) fusions of the promoter region of the gene of interest (*goi*) to *lacZ*. The DNA from the vector is shown in black, and the genetic elements are in light grey. The DNA and the genetic elements from the inserted DNA are in dark grey. Transcriptional terminators are symbolized by hair-pins and the other genetic elements (Shine-Delgarno: Ribosome binding sites, fMet: Initiator codon) are shown as boxes.

3.2.2. Green fluorescent protein

Originating from the Pacific Northwest jellyfish *Aequorea victoria*, GFP is a widely used reporter possessing a wealth of applications. It can be used as a reporter protein in all cell types (see Fig. 4), including bacteria (Chalfie et al., 1994; Christensen et al., 1996) and yeast (Kahana et al., 1995; Niedenthal et al., 1996). Besides the usage as a reporter of promoter activity, GFP is used as fusion tags to observe protein localization within living cells.

GFP is a small bioluminescent protein of 238 amino acid residues. Its chromophore is spontaneously generated by cyclization of Ser65-Tyr66-Gly67 and 1,2-dehydrogenation of the tyrosine. With the exception of oxygen needed for the activation of the chromophore formation, GFP does not require substrates or other cofactors to generate green light (Chalfie et al., 1994; Heim et al., 1994; Tsien, 1998).

The emission and absorption spectra of wild type GFP have been optimized for fluorescence-activated cell sorting (FACS; cf. Section 3.3) and fluorescent microscopes (Cormack et al., 1996; cf. Section 3.4) and alternative proteins with emission peaks ranging from blue to yellow have been developed (Heim et al., 1994; Heim and Tsien, 1996; Tsien, 1989). Similarly, the folding efficiency and thermostability of the relatively unstable wild type GFP have been optimized (Crameri et al., 1996; Patterson et al., 1997; Yang et al., 1996). Besides taking into consideration which GFP mutant will be optimal for the experimental setup, the codon usage of the heterological reporter gene is also an important factor to include. Fusion to a highly expressed protein could be a solution to defeat slow translation initiation (Veening et al., 2004).

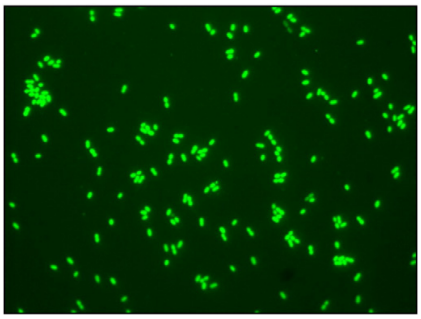

Fig. 4. Epi-fluorescence image of a *Lactococcus lactis* strain emitting green fluorescence. The gene encoding GFP is expressed from a *L. lactis* promoter and the construct is present in a single-copy on the chromosome.

GFP has the ability to serve as a pH probe in aqueous solutions and intracellular compartments in living cells, as different pH induces changes in the fluorescence intensity (Kneen et al., 1998). While many GFP mutants display good pH responsiveness, the excitation spectrum of wild type GFP is essentially unaltered between pH 5.5 and 10 (Tsien, 1989; Ward, 1981; Ward and Bokman, 1982). GFP mutants have been constructed with the ability to display changes in the emission pattern dependent on the pH, thereby making the analysis independent of the protein concentration (Hanson et al., 2002; Miesenböck et al., 1998).

Similarly, whole-cell biosensors based on GFP fusion with stress-regulated promoters can be used for sensing toxicity (Norman et al., 2006; Sørensen et al., 2006), and fusions with growth dependent and stress dependent genes (Brauer et al., 2008) may yield a more sophisticated understanding of single-cell behavior.

3.2.3. Protein–Protein interactions

Several reporter systems for protein–protein interactions exist and these may be based on colorimetric reporters (*lacZ*), auxotrophic markers or fluorescent bioluminescent molecules. The yeast two-hybrid system, which is still intensively used for protein–protein investigations, takes advantage of the modular nature of many eukaryote gene regulators that consist of a DNA binding domain (DBD) and a DNA activating domain (AD). A protein of interest, the bait, is translationally fused to the DNA binding domain (DBD) of a gene regulator, typically the DBD of GAL4. Another protein of interest, the prey, is fused to the DNA activating domain typically that of GAL4.

The DNA binding domain (DBD) binds to upstream activating sequences (UAS) of GAL4 on the chromosome. If the two proteins, bait and prey, interact, this interaction will bring the AD into proximity of the promoter, thereby recruiting DNA polymerase II and exert transcription of the reporter gene downstream of the promoter (Chien et al., 1991; Fields and Song, 1989). The two-hybrid system was originally developed for the yeast *S. cerevisiae*, but this type of system is also available for prokaryote and mammalian cell lines (Fearon et al., 1992; Joung et al., 2000). This two-hybrid system screens for interactions between cytosolic proteins, but modifications of the system have been developed that allow for screening for interactions between various biomolecules. Protein–DNA interactions can be analyzed in the one-hybrid screen (Li and Herskowitz, 1993), protein–RNA interactions in the three-hybrid screen (Zhang et al., 1999) and a system similar to the yeast two-hybrid, the split ubiquitin system, which allows for analysis of interactions between membrane molecules has also been developed (Stagljar et al., 1998). Also split *lacZ* (Wigley et al., 2001) and split-GFP-based bimolecular screens exist (Hu et al., 2002).

3.2.4. Protein solubility and stability

Translational fusions to GFP can be used for measuring protein solubility. The fluorescence signal of *E. coli* strains expressing proteins fused GFP have proved to be related to protein folding (Waldo et al., 1999). The proteins may be fused to the N-terminal of GFP or as internal fusions within the GFP reading frame (Cabantous et al., 2005). Fusion to GFP may affect solubility of target proteins. However, a modified fusion method where a 15 amino acid GFP fragment (GFP-11) is fused to the target protein allows, due to the small size of the GFP fragment, for minimal effects on target protein solubility (Cabantous et al., 2005). The remaining major part of GFP (GFP1-10) is expressed separately. The two GFP fragments associate spontaneously and form fluorescent GFP. However, if the target protein fails to fold correctly, this prevents accessibility to GFP-11 and thus prevents a fluorescent signal. Protein stability can also be quantified using reporter constructs. By transcriptional fusions of *Discosoma sp*. Red (DsRed, a red fluorescence marker) and GFP tagged to a protein of interest, the GFP/DsRed signal ratio has been shown to report on protein turn over in the cell (Yen et al., 2008).

3.2.5. Förster Resonance Energy Transfer or Fluorescence Resonance Energy Transfer

Förster Resonance Energy Transfer or Fluorescence Resonance Energy Transfer (FRET) relies on nonradiative energy transfer between two chromophores and is another popular means for measuring protein–protein interactions and protein conformational changes. Two proteins are tagged with a donor- and an acceptor chromophore, respectively. The most popular donor–acceptor pair is cyan fluorescent protein, CFP (donor) and GFP (acceptor) each of them translationally fused to one of the proteins of interest. If the two proteins of interest are dissociated, donor emission is observed upon donor excitation, while in case the proteins associate (i.e. are in the proximity of 1–10 nm), acceptor (and donor) emission is observed upon donor excitation (Jares-Erijman and Jovin, 2003, 2006). When protein conformational changes are investigated, one protein is labeled with a donor at one end and an acceptor at the other end. Conformational changes in the protein are then observed as changes in the FRET signal. Like for the two-hybrid system, versions of FRET have been developed for interactions of non-cytosolic proteins such as receptor–ligand interactions as well as for protein function (Prinz et al., 2008).

A similar technique, Bioluminescence Resonance Energy Transfer (BRET), utilizes the bioluminescent luciferase instead of a fluorescent donor molecule (Pfleger and Eidne, 2006). BRET, unlike FRET, does not require a source of external illumination, but the ATP hydrolytic activity of luciferase increases the risk of secondary effects on cell physiology when using BRET.

3.3. Microscopy

Microscopic methods provide powerful tools to investigate microbial heterogeneity with single-cell resolution. Both fluorescence and light microscopy coupled to image analysis are very useful for studies of individual cells in fermenting populations and can greatly contribute to the knowledge about microbial population heterogeneity. Several stains originally developed and applied in microscopy are now primarily used in flow cytometry (cf. Section 3.4). Despite the advantages of using a high-throughput method as flow cytometry, several stains and uptake of radiolabeled substrates are only or best analyzed by fluorescence or light microscopy such as microautoradiography (MAR).

3.3.1. Fluorescence microscopy

Fluorescence-based microscopy methods are useful for a wide diversity of applications ranging from industrial to environmental microbiology (Joux and Lebaron, 2000). Stains that allow demonstration and quantification of nucleic acids, proteins, lipids or that stain polyester or polyphosphate inclusion bodies, as well as assessing membrane integrity are commercially available (Brehm-Stecher and Johnson, 2004; Shapiro, 2000). Multiple fluorescent stains may be used simultaneously, allowing collection of multiple parameters per cell and thus giving more accurate evaluation of cell physiology (Brehm-Stecher and Johnson, 2004; Joux and Lebaron, 2000). A demonstrative multicolor fluorescence test with seven fluorochromes applied on a single biological specimen has been presented by Adav et al. (2010). The main attributes of fluorescence as a tool in microscopy are its specificity, sensitivity, temporal and spatial resolution (Maukonen et al., 2003), and problems that might be encountered include fading, photo-bleaching of the fluorochrome, and fluorescence-quenching (Maukonen et al., 2003).

Epifluorescence microscopy and confocal scanning laser microscopy (CSLM) are used for studying specimens that fluoresce. The CSLM technique provides detailed, non-destructive examination of thick microbial samples by controlling the depth of the scanning field and the collection of serial optical sections. The latter allows for the generation of three-dimensional images and thus the spatial analysis of a microbial population. It also provides fluorescent scanning of substances bound with different fluorochromes that are excited using light at different wavelengths (Adav et al., 2010; Joux and Lebaron, 2000; Maukonen et al., 2003).

Image analysis has become a valuable accessory for rapid quantification purposes and handling of collected data in a wide diversity of applications. It reduces subjectivity caused by manual operation and allows automation. In some fields such as fermentation technology, image analysis is essential for characterizing the physiological state of the culture (Carneiro et al., 2009).

Several programs for quantification purposes are readily available. IMAGEJ is a general program that is useful in many research fields, though several other specific programs (e.g. DAIME developed by Daims et al., 2006) were developed according to specific needs (Neu et al., 2010). Staining techniques allow determination of individual cell viability, activity, surface properties and internal storage compounds and can contribute to knowledge about population heterogeneity. Examples of dye and probe based techniques are presented below and summarized in Table 1.

3.3.1.1. Detection of cell viability. Cellular viability is one of the most relevant physiological parameters to be assessed. Membrane integrity is the most commonly used indicator of viability and is based on the capacity of cells with intact membranes to exclude fluorescent dye. Because of the fact that nucleic acid stains exhibit a large enhancement of fluorescence upon binding, and since those molecules are present in the cells at high concentrations, nucleic acid stains have been widely used for membrane integrity assays (Carneiro et al., 2009). Live/Dead staining is a method commonly used for the assessment of membrane integrity (Boulos et al., 1999). It is based on two fluorescent stains: SYTO 9, which penetrates all bacterial as well as yeast cells, and propidium iodide stain (PI), which penetrates only cells with compromised membranes, causing a reduction in the SYTO 9 fluorescence when both dyes are present due to higher affinity for nucleic acids. When both dyes are used in combination, viable cells fluorescence in green whereas dead or damaged cells fluoresce in red. Live/Dead staining kits are commercially available (BacLightTM) and have already been successfully applied to a variety of fermentation studies (Carneiro et al., 2009; Corich et al., 2004; Zotta et al., 2009) where pure cultures were examined, as well as for investigation of more complex, environmental samples (Freese et al., 2006; Hao et al., 2009; Larsen et al., 2008).

Although PI seems be the most stringent indicator for membrane integrity, its utility as universal indicator for microbial cell death should not be overestimated. It was observed by Shi et al. (2007), that for some species, viable cells from the early exponential phase can incorporate PI and give a false positive signal. Another problem

Table 1
Examples of dye and probe based techniques for measurement of cell properties.

Properties	Indicator	Techniques
Cell viability	Membrane integrity	SYTO 9
		Propidium iodide (PI)
		Ethidium monoazide (EMA)
		Propidium monoazide (PMA)
		DNA stains
		Carboxyfluorescein diacetate (CFDA)
	+ Detection of membrane carbonylated proteins	Alexa Fluor hydrazide fluorescent dyes
Cell activity		
Respiratory activity	Reduction of formazan crystals by the cell electron transport system	5-cyano-2,3-ditolyl tetrazolium chloride (CTC)
	Reductase activity in cell electron transport system	Redox Sensor Green (RSG)
Metabolic activity	Ribosome concentration	Fluorescent probes used in Fluorescence In Situ Hybridization (FISH)
Microbial surface properties	Cell surface hydrophobicity (CSH)	Fluorescent polystyrene microspheres designed for Microsphere Adhesion to Cells (MAC)
	Cell surface polysccharides	Fluorescent lectin
	Amyloidic fibrils	Fluorescent antibodies
	Various exoenzymes	Fluorescent dye-labeled substrates
Internal storage compounds	Polyhydroalkanoates (PHA)	Nile blue

reported by several authors (Lisle et al., 1999; Stewart and Franklin, 2008) concerned the formation of a sub-population of cells containing simultaneously both SYTO9 and PI inside the cells. These artifacts appear to be live and dead simultaneously. As an explanation, it was proposed that in cells with minimal membrane damage, PI has limited access to the cytoplasm, allowing it to accumulate to concentrations high enough to be visualized, but too low to totally quench SYTO 9 (Millard and Roth, 1997).

Another fluorescent stain used to assess cell viability on the basis of membrane integrity is ethidium monoazide (EMA). This dye penetrates only cells with compromised membranes (Riedy et al., 1991). The technique was successfully applied by Regan et al. (2003) on a pure culture of nitrite oxidizers to quantify the fraction of dead cells in a population. However, Nocker et al. (2006) compared EMA to propidium monoazide (PMA), which is another membrane impermeant dye used for viability assessment, and proved that PMA exhibits better selectivity towards dead cells staining than EMA.

Poor staining of dead cells whose DNA is degraded can be a common limitation for all DNA stains used to assess membrane integrity (Joux and Lebaron, 2000). To overcome this problem another approach utilizing Alexa Fluor hydrazide fluorescent dyes for the detection of dead cells was developed. The membranes of living cells are impermeable for this chemical, which can be used for the detection of carbonylated proteins (Ahn et al., 1987). The quantity of these proteins inside the cells increases after cells are exposed to stresses or when they are in a stationary phase of growth (Saint-Ruf et al., 2010). Using this method Saint-Ruf et al. (2010) performed precise quantification of the fraction of dead cells in stressed and non-stressed *E. coli* populations. Additionally, carboxyfluorescein diacetate (CFDA) can be used as an indicator of microbial viability as cell-membrane integrity is required for retention of fluorescent staining product inside the cell (Joux and Lebaron, 2000; Miyanaga et al., 2007).

3.3.1.2. Detection of cell activity. The compound 5-cyano-2,3-ditolyl tetrazolium chloride (CTC) is colorless and membrane-permeable and is capable of producing insoluble, red-fluorescing formazan crystals in the cell when it is reduced by the electron transport system of bacterial cells. This method is commonly applied to determine respiratory activity and viability of single bacteria in cultures as well as in many different complex microbial systems (Rodriguez et al., 1992). It was used by many research groups for determination of active bacteria both in pure cultures (Créach et al., 2003; Lisle et al., 1999) as well as in environmental samples (Freese et al., 2006; Nielsen et al., 2003a; Pyle et al., 1995). Unfortunately, as reported by several researchers, CTC results are not always consistent with results obtained with other methods used to determine the fraction of active cells in a population (Créach et al., 2003; Nielsen et al., 2003a) and sometimes tend to underestimate the real number of active cells. It has been suggested that CTC is toxic to the cells and targets only the most active cells (Joux and Lebaron, 2000; Sherr et al., 1999; Stewart and Franklin, 2008). Additionally, problems associated with this method concern the lack of a standardized protocol for using CTC and the fact that results can be biased by CTC-media composition (Pyle et al., 1995; Smith and McFeters, 1997).

Some authors point out the potential applicability of Fluorescence In Situ Hybridization (FISH) for assessment of bacterial metabolic activity. This approach is driven by the assumption that high ribosome content in the cell would indicate physiological activity of this cell, as the FISH signal intensity is a reflection of the ribosome concentration within the cell (Poulsen et al., 1993). Thus changes of FISH signal intensity could be quantified and would reflect fluctuations of the cell metabolic activity (Daims and Wagner, 2007; Wagner et al., 2003). However, the authors point out that this assumption is not true for all bacterial species as it was proven that the ribosome content in some bacteria is not correlated to their metabolic activity. To improve the method the use of oligonucleotide probes that bind to the spacer regions between the rRNA genes is recommended (Oerther et al., 2000; Schmid et al., 2001; Wagner et al., 2003). Indeed, the synthesis of pre-rRNA inside the cell stops as bacterial growth ceases. Consequently, the pre-rRNA pool in non-growing cells is depleted due to the continuous ribosome assembly process, being though restored when growth resumes (Cangelosi and Brabant, 1997). This type of FISH would then more precisely reflect bacterial activity.

3.3.1.3. Characterization of microbial surface properties. Cell surface hydrophobicity (CSH) of bacteria could potentially be used for the characterization of heterogeneous microbial populations as the metabolic state may determine bacterial surface properties. Microsphere Adhesion to Cells (MAC) was developed by Zita and Hermansson (1997) to study CSH. It includes the use of fluorescent polystyrene microspheres with a diameter of 0.1 μm (or smaller) and defined hydrophobicity, which attach to the cell surface. The method was applied by Nielsen et al. (2001) to determine CSH of specific bacteria in activated sludge and it might also be applicable to bacterial cultures. However, Heard et al. (2009) point out that interactions between microspheres and cell surfaces are not simple as they are controlled not only by hydrophobic effects but also by other types of forces, e.g. electrostatic interactions. The authors point out that surface chemistry of microspheres should be carefully considered before the experiment.

3.3.1.4. Detection of internal storage compounds. Many bacteria are able to accumulate storage compounds, which can serve as a reserve of carbon and energy sources during periods of starvation and growth

e.g. polyhydroxyalkanoates (PHA) (Waltermann and Steinbuchel, 2005). PHA content in single cells can be visualized with a Nile blue stain using the method developed by Ostle and Holt (Ostle and Holt, 1982). As the fluorescence intensity is proportional to the PHA content of the cells (Page and Tenove, 1996), the PHA amount can be quantified and the results can be used for investigation of bacterial population activity. Kragelund et al. (2005) determined the distribution of PHA content among a *Meganema perideroedes* population in activated sludge and used it as activity measure. Although the application of this method was reported only for the activated sludge sample, it presents the potential to be applied to industrial fermenting populations.

3.3.1.5. Other fluorescent stains. A range of other stains is available for studies of single cell characteristics by microscopy and in some cases flow cytometry. These have, however, not yet been used to study heterogeneity in fermentations but many may have a great potential. These include the novel fluorogenic dye Redox Sensor Green, which can be used for in situ detection of actively respiring cells (Kalyuzhnaya et al., 2008) as well as staining of different surface components such as polysaccharides (Neu et al., 2010), amyloidic fibrils (Larsen et al., 2007) and various exoenzymes (e.g. Kloeke and Geesey, 1999; Kragelund et al., 2005).

3.3.1.6. Quantum dots. Although not yet employed for examination of fermenting population heterogeneity, Quantum Dots (QDs) present a great potential as alternative to fluorophores to be utilized by fluorescent microscopy in this type of studies. QDs are nanocrystalline objects with diameter from 2 to 10 nm having broad absorption and narrow emission spectra, which are dependent on composition and size of the QD. They are available in a very broad range of well separated colors. Their unique features are broadly described in literature. QDs are up to 20 times brighter than organic dyes and are resistant to bleaching. Long fluorescence lifetime of QD enables separation of their signal from background fluorescence and unique spectral features allow detection and tracking of multiple QDs signals (Alivisatos, 1996; Michalet et al., 2005; Mazumder et al., 2009; Walling et al., 2009; Bae et al., 2010; Fu et al., 2009). In prokaryotic systems QDs have been applied mainly for external (cell surface) labeling (e.g. Kloepfer et al., 2003; Zhao et al., 2004; Chalmers et al., 2006; Bae et al., 2010; Mazumder et al., 2010). Metabolism-specific labeling of bacteria was performed with the use of QD-adenine conjugates (Kloepfer et al., 2005) as well as QDs functionalized with citrate, isocitrate, succinate, or malate (Hirschey et al., 2006). Lately, Wu et al. (2010) applied FISH with QD-based Molecular Beacon to specifically detect β-lactamase genes located in a recombinant *E. coli* plasmid. Although very promising, QDs possess several limitations, which have to be considered prior to application. They have a larger size than the conventional fluorophores, they sometimes exhibit an on/off behavior resulting in "blinking" of the signal (Walling et al., 2009; Mazumder et al., 2010) and, most importantly, QDs can be toxic towards bacteria due to leakage of heavy metals as well as formation of toxic compounds caused by QDs surface oxidation (Kloepfer et al., 2005; Schneider et al., 2009).

3.3.2. Conventional light microscopy

Light microscopy alone or in combination with epifluorescence microscopy can be applied to determine the substrate uptake by individual bacteria by microautoradiography (MAR). The method is based on the observation that the radiolabeled substrate taken up by individual cells can be visualized with a radiation-sensitive silver halide emulsion, which is placed over the radiolabeled bacteria and subsequently processed by standard photographic procedures (Carman, 1993; Nielsen and Nielsen, 2005). It is a powerful tool used first in microbial ecology in the 1960s (Brock and Brock, 1966, 1968) and can be used to detect cell viability, enumeration of bacteria capable of consuming specific organic substrates, studies of autotrophic activity, uptake of orthophosphate and potential use of various electron acceptors. The method is usually applied in a semi-quantitative way so bacteria in a population are determined to being active or non-active. However, it is possible to perform quantitative MAR (q-MAR) so the distribution of activities of individual cells in a certain population can be analyzed (Nielsen et al., 2003b). MAR can be combined with different fluorescent stains simultaneously giving a more accurate evaluation of cell physiology. Examples are MAR combined with CTC and FISH (Nielsen et al., 2003a).

In mixed microbial systems with several species, MAR can also be combined with FISH (MAR-FISH) for identification of the different bacteria so aspects of the ecophysiology of the specific species can be investigated (Lee et al., 1999; Nielsen and Nielsen, 2005).

The radiotracers used are typically the soft beta-emitters like ^{3}H, ^{14}C and ^{33}P. Common for all is that the silver grains on top and near the labeled bacteria can be visualized by bright field or phase contrast microscopy. Low energy emitters give the highest resolution. Tritium, for example, has a resolution of approximately 0.5 μm while that for ^{14}C and ^{33}P is 2–3 μm. A detailed description of general procedures for MAR can be found in several publications (e.g. Carman, 1993; Nielsen and Nielsen, 2005).

3.4. Flow cytometry

Flow cytometry (FCM) is a tool that counts, sorts and examines objects in suspension such as bacteria or yeast. It is a robust technique that relies on the properties of light scattering, excitation and emission to measure a variety of properties of single cells. The ability of FCM to measure the properties of single cells allows the study of phenotypic diversity of individual microorganisms (Brehm-Stecher and Johnson, 2004).

The extensive variety of cellular parameters, that can be studied simultaneously, and the facility to acquire information on how such parameters spread in a cell population is of core importance to the method. When cells pass through a light source, unique electronic and optical parameters are measured based on groupings, succession and/or ratios of selected parameters. The information obtained through measurement of the selected parameters can then be linked with different cell properties and components (Mandy et al. 1995).

Different methods have been developed in order to study diverse cell properties such as size, intracellular pH and membrane potential that can indicate diverse cellular characteristics such as the levels of cellular components as including DNA, calcium, protein and surface receptors.

Measuring properties of single cells within an entire population can provide a more accurate and descriptive representation of the population than average values attained from traditional techniques (Rieseberg et al., 2001). Indeed, due to the possibility of measuring distributed properties in cell cultivations, flow cytometry is a useful tool in the study of heterogeneity in microbial populations (Davey and Winson, 2003), and may provide valuable understanding for bioprocess design and control (Diáz et al., 2010). It has for example been used for monitoring dynamic changes in yeast gene expression (Mateus and Avery, 2000), for control of biomass concentration (Kacmar et al., 2006), for quantification of horizontal gene transfer in bacterial populations (Sørensen et al., 2003), as well as for studying heterogeneity of stress gene expression (Atfield et al., 2001). The role of flow cytometry in molecular biology, with reference to gene reporter systems, has been reviewed by Davey and Winson (2003). Also several different industrial applications of flow cytometry have been reviewed by Diáz et al. (2010).

3.4.1. Flow cytometer: principles

Within a flow cytometer, single cells are hydrodynamically focused in a fluid stream within a carrier fluid known as the sheath

fluid. The control of the system's pressure and the orifice size allows the establishment of a cellular laminar flow regime. This hydrodynamic focusing of the slower moving sample at the center of a rushed flow stream creates a high speed single cell flow that is intercepted by a light source (usually a laser beam) for the interrogation of particles' properties. The light scattered or emitted by the cells or by cell-associated fluorophores can be isolated and optically separated by collection optics, mirrors and filters. If the angle of deviation of light traveling through the cell flow is small, it will be detected as forward scatter (FSC). Conversely if the angle of light deviation is big it will be detected as side scatter (SSC). After collection with appropriate wavelength filters, fluorescence is identified on a big angle detector. This light will ultimately trigger a photomultiplier tube (PMT) that augments the signal and finally converts it into a digital signal. Thus FCM does not usually present real images of the bacteria as fluorescence microscopy (cf. Section 3.3). Instead digital data is produced and presented as dual-parameter plots or a single-parameter histogram (Bergquist et al., 2009; Czechowska et al., 2008; Davey and Kell, 1996; Link et al., 2007) (see Fig. 5).

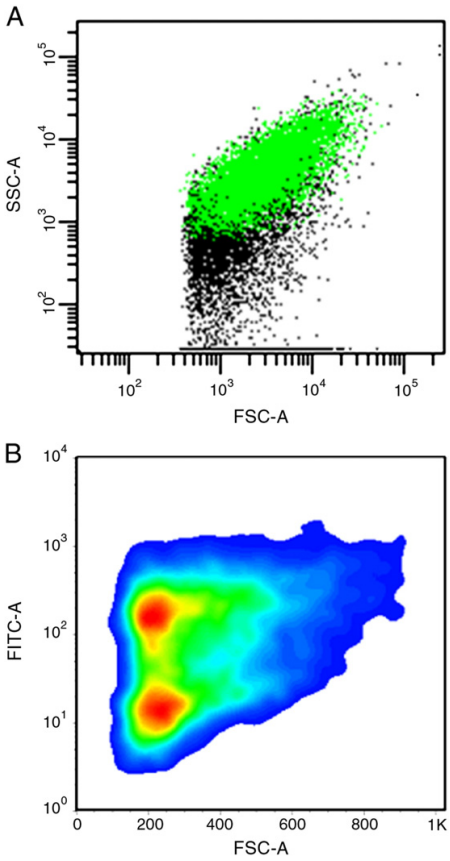

Fig. 5. Distributed measurements of cell properties obtained by flow cytometry analysis. (A) Bivariate distribution (Side scatter vs. Forward scatter) of GFP expressing *E. coli* (in green) within a mixed bacterial population; (B) Bivariate distribution (GFP Fluorescence vs. Forward scatter) of a *S. cerevisiae* growth rate reporter strain expressing GFP during exponential growth on glucose.

3.4.2. Measurable cell properties

The cell properties measured by FCM can be classified into intrinsic and extrinsic depending on the method of measurement. Intrinsic properties such as size or membrane composition can be studied without the need to label the cells, in the FSC and SSC channels. Moreover, the electronic volume measurement is commonly used for cell size determination (Shapiro, 2004). Extrinsic parameters, on the other hand, normally use fluorescent stains or fluorescence labeled probes to study microorganisms' characteristics and components such as membrane integrity or potential (Mandy et al., 1995). Given that discrimination of dissimilar cell types and background is attained using fluorescent labels, various fluorescent dyes and labels, also used for microscopy studies (cf. Section 3.3.1), are available for targeting specific biological materials (e.g. nucleic acids) or to signal biological activities (e.g. enzyme activities or membrane potential), leading to a deeper understanding of physiological and metabolic functions in bacteria (Ishii et al., 2010). Today there is a wide availability of probes and labels available, and the most adequate should be chosen considering the specific strain and cultivation conditions which will be studied (Shapiro, 2004).

The measurement of the three parameters (FSC, SSC and fluorescence) is processed in real-time with help of specific software packages supplied with the commercial instrument. Most of the FCM software allows the operator to define areas of specific interest in a procedure called “gating”. Gating works by digitally filtering the FSC and SSC and fluorescence signals, targeting a subset of results. For single-organisms gating can discriminate cells of similar properties from a population or identify bacteria within a complex sample (Bahl et al., 2004; Mandy et al., 1995; Hammes and Egli, 2010).

Due to different cell properties as size and shape, it is possible to discriminate cells based solely on size and light scatter characteristics and consequently, for example, differentiate bacteria from yeast (Veal et al., 2000).

3.4.3. Fluorescence-activated cell sorting

An important extension of FCM is the ability of some flow cytometers to physically sort cells according to their light scattering and fluorescent properties, this feature is commonly known as fluorescence-activated cell sorting (FACS) — this term, coined by the Becton-Dickinson company, is now generically used by the scientific community (Tracy et al., 2010). Depending on the instrument used, there are different methods to capture a cell of interest. Nevertheless, the sorting is always based on one or multiple combinations of user defined cell characteristics. Practically speaking, FACS allows the user to separate interesting sub populations of cells from complex mixtures for further analysis on the basis of user defined limits to FSC, SSC and fluorescence obtained in a given experiment (Ishii et al., 2010).

In sum, this automated flow analysis of cells has the combined advantages of allowing multi-parameter data acquisition and multivariate data analysis, high throughput and high-speed analysis and, as mentioned, the ability to sort cells according to any combination of parameters processed. Thus, FCM and FACS have inestimable benefits for microbiology studies.

Detailed description of FCM and FACS principles are reviewed elsewhere (Brehm-Stecher and Johnson 2004, Link et al. 2007, Bergquist et al., 2009; Davey and Kell, 1996; Shapiro, 2004).

Combinations of FCM and FACS techniques with fluorescent staining or labeling are readily used to shed light onto pertinent areas of microbiological interest, namely by allowing the study of three crucial aspects of the microorganism: the physiological state of the microorganism, the transcription and protein biosynthesis, and the measurements of structural or physical parameters of a cell (Tracy et al., 2008, 2010). Thus, the techniques can be applied to study the types of input that individual cell differences have and how these contribute to heterogeneity within a microbial population.

3.4.4. Physiological state of microorganisms

The study of the physiological state of a cell has been redefined by the application of FCM as a tool. The term microbial "viability" gained another dimension upon the disclosure that microorganisms are not just alive or dead, but also have a range of transitional states (Nebe-von-Caron et al., 2000). Existing fluorescent nucleic acid stains are either able to permeate the membrane or not and the combination of both types of stains, for example SYTO dyes (green fluorescence) and PI (red fluorescence), is used in membrane integrity assays (Allegra et al., 2008; Ben-Amor et al., 2005; Berney et al., 2007). Furthermore, the cellular membrane potential is used as a test of viability. Depending on the membrane potential and whether the dye used is cationic or anionic, the cell emits (or not) fluorescent signals with different intensities (Nielsen and Sjøholm, 2009; Novo et al., 1999). Measuring the changes in membrane composition directly related with changes in the cellular physiological state is also possible by using the stain 1,6-Dephenyl-1,3, 5-hexatriene (DPH) (Müller et al., 2000). Testing for enzyme activity is generally done by allowing a membrane permeable nonfluorescent substrate to be taken up by the cell and be metabolized into a preferably impermeable fluorescent substrate. Based on this principle, respiring and non-respiring cells can be distinguished by the measurement of dehydrogenase activity with, for example, CTC (Falcioni et al., 2008; Gasol and Arstegui, 2007) although this method does not always present consistent results (Caro et al., 2007). Another example of enzyme activity that can be studied with similar practices is esterase activity by staining with, for example, CFDA (Hoefel et al., 2003). Membrane pump activity can also be assessed by loading different dyes such as PI or ethidium bromide, into the cell and measuring fluorescence reduction (Bunthof et al., 1999; Nielsen and Nielsen, 2005). Another test of viability is the measurement of internal pH and its variation with culture change, using a probe whose fluorescent characteristics correlate with changes in pH (Cronin and Wilkinson, 2008; Rault et al., 2009).

3.4.5. Transcription and protein biosynthesis

As mentioned before, in transcription and protein biosynthesis studies two main strategies are commonly employed: one approach is to use genetic constructs with reporter genes expressing GFP and another is the development of FISH techniques that successfully have been adapted to FCM. Both strategies enable researchers to study specific functions in microorganisms at a single-cell level, in real-time. Induction of GFP in reporter bacteria was used as a viability test in a study of the effect of pasteurization (Gunasekera et al., 2002). Reporter genes expressing GFP in bacterial biosensors can be designed specifically as micro-analytical devices that are detected and quantified by FCM. Examples of this approach are the detection of quorum sensing (Burmølle et al., 2003, 2005), the genotoxic complex mitocin C metabolized by *Streptomyces* (Norman et al., 2006) and horizontal gene transfer (Sørensen et al., 2003).

The combination of FISH and FCM is commonly referred as FLOW-FISH. Here, fluorescence labeled oligonucleotides based upon rRNA sequences are inserted in cells and annealed to complementary RNA sequences. This method is primarily used to quantify and screen mixed populations targeting distinct species (Fornasari et al., 2008; Jen et al., 2007; Kalyuzhnaya et al., 2006). Different strategies are being constantly developed to boost the fluorescent signal. In theory this approach can be used to target cells expressing specific mRNAs within a given microbial community, however the few reports available document that FLOW-FISH has so far only been successful in laboratory cultures (Jen et al., 2007; Kalyuzhnaya et al., 2006).

3.4.6. Structural or physical parameters of a cell

The measurements of structural or physical parameters of a cell are mostly based on FSC and SSC readings, as well as electronic volume measurements, and are used for analysis of the shape and size of the cell, Gram staining characteristics, cell surface and antibiotic-binding sites (Forster et al., 2002; Papadimitriou et al., 2007; Tracy et al., 2008).

In conclusion, it should be stated that some bias can be induced in FCM analysis, especially related to difficulties associated with the establishment of accurate and validated reading gates for each sample and with the danger of signal quenching when using multiple probes simultaneously. Although efforts in developing automated gating algorithms have been published (e.g. Lo et al., 2008; Pyne et al., 2009), the development of data analysis tools (software) has not followed the technological progress of the flow cytometer (hardware), as consequently, the still frequent need for manual analysis of individual samples is indicated as a limiting aspect of the flow cytometry technology (Lo et al., 2008). Multivariate and artificial intelligence approaches to flow cytometric data analysis, including cluster analysis and artificial neural networks, have been reviewed elsewhere (Bashashati and Brinkmand, 2009; Davey and Davey, 2010). FCM methods are highly versatile and applicable, and the speed of analysis and the multivariate data sets produced are attractive for developing models for distributed properties. Efforts to get consensus and broader acceptance of microbial cytometry method standardization efforts for procedures, instruments, as well as data analysis and presentation, are currently under development (Nebe-von-Caron, 2009; Lee et al., 2008).

3.5. Raman spectroscopy

The use of spectroscopy for the on-line monitoring of bioprocesses is becoming more common, due to its non-destructive nature, and the ability of providing information at molecular level without the use of stains or radioactive labels (Krafft et al., 2009). In particular Raman spectroscopy has been successfully applied to bioprocess analysis (Lee et al., 2004; Ulber et al., 2003). This technique is based on shifted wavelength scattering of light resulting from inelastic collisions of photons with the molecules, upon excitation of the sample with monochromatic light (Ulber et al., 2003). By single-cell analysis of bacteria, it has been proven that cells from one culture sample, however morphologically similar, show different Raman spectra (Schuster et al., 2000a,b) due to differences in the intracellular content.

Confocal Raman microspectroscopy has been used to investigate the spatial intra- and inter-cellular heterogeneity in genetically homogeneous microbial cultures in a rapid and non-destructive way (Hermelink et al., 2009). The spectra obtained for microbial cells reflect their biochemical and structural composition. Indeed, Raman imaging can be used to visualize the distribution of cellular components, such as proteins, DNA, RNA, intracellular lipid vesicles, mitochondria, and chromosomes, within a single cell (Krafft et al., 2009). The spatial heterogeneity of eukaryotic cells such as yeasts resulting from inner compartmentalization was for example studied, by means of Raman microspectroscopy, by recording several spectra at different locations inside the same yeast cell (Hermelink et al., 2009).

In conjugation with infrared absorption spectroscopy, which yields information at microbial colonies level rather than at single-cell level, Raman spectroscopy has been used to identify and characterize heterogeneity in both microbial micropopulations cultured for short periods of time (6–10 h) and single cells, with minimal sample handling and no need for dyes or contrasting agents (Choo-Smith et al., 2001; Schuster et al., 2000a).

Intracellular heterogeneity reflecting the predominance of certain components (e.g. proteins, lipids) and cell compartmentalization has been reported based on Raman spectra collected at different spatial positions within a single-cell (Rösch et al., 2005a,b). This allows for monitoring of microbial growth in cases where different components show significant changes in quantity and localization within the cell,

at a certain growth phase (e.g. the formation of poly-β-hydroxybutyric acid in *Bacillus cereus* when a carbon supply is accessible (Hermelink et al., 2009)).

Raman spectroscopy offers, therefore, the possibility of monitoring cell heterogeneity in microbial populations by chemical and structural imaging. It can, consequently, contribute to a deeper understanding of the role of for example cell aging and cell cycle in generating a phenotypically heterogeneous cell population (Rösch et al., 2005b).

4. Designing dynamic experiments at different scales

As previously referred (cf. Section 1), heterogeneous microbial populations have been observed to develop in large scale bioreactors as a result of experiencing a different extracellular microenvironment as they circulate in the bioreactor. In order to experimentally study the cell-environment interplay and, consequently, better understand the origins of cell heterogeneity, it is necessary to be able to experimentally simulate such environmental gradients observed in large-scale reactors. In this section, we present strategies to perform such studies including the use of scale down reactors (cf. Section 4.2), pulse experiments (Section 4.3), and microbioreactors (cf. Section 4.4), as well as discuss the importance of sampling strategies and bioprocess control (cf. Section 4.5).

4.1. Heterogeneities in large scale reactors

It may not be practically possible to sustain homogeneity at a large scale, since this may demand an unrealistic amount of energy input (Hermelink et al., 2009). Therefore, significant gradients of dissolved oxygen, substrates, pH and dissolved carbon dioxide are often encountered. For example, oxygen gradients are frequently found in large scale bioreactors due to transport limitations when the oxygen consumption rate exceeds the oxygen transport rate, which is a situation commonly seen for high cell density cultures.

Furthermore, in some process modes, for example in fed-batch cultivations, it is inherently impossible to avoid substrate gradients, since a highly concentrated substrate is fed in a narrow zone. For example, a fed-batch cultivation of *Saccharomyces cerevisiae* may involve addition of glucose at concentrations as high as 600 g/L to avoid a significant increase in volume (Larsson et al., 1996). Consequently, the microbial cells experience rapid changes in environmental conditions as they circulate throughout the reactor, which might pose stress on the cells and affect their metabolism.

A heterogeneous environment has previously been shown to generate cell population heterogeneity in *E. coli* fed-batch cultivation in a large scale bioreactor (22 m^3) (Larsson et al., 1996). Enfors and coworkers demonstrated that the transcription level of four genes related to low-oxygen or high-glucose concentration varied in cells that were sampled from different zones of the reactor with short or long distance to the glucose feed zone (top) and the oxygen inlet point (bottom), thus illustrating the development of sub-populations within the bioreactor as a response to gradients (Enfors et al., 2001).

Interestingly, *E. coli* cell viability was higher at large scale than in bench-scale; in fact, a close to 100% viability was reported after 40 h of cultivation compared to only 85% in bench-scale, in contrast to the biomass yield, which was higher at bench-scale (Enfors et al., 2001). This demonstrates that population heterogeneity may have both positive and negative consequences, and hence that a certain level of heterogeneity may be beneficial in order to better cope with the changing conditions in large reactors.

4.2. Scale down reactors

To better predict the results that can be achieved in large scale, several types of scale down reactors (SDRs) have been developed and reported to better represent large scale conditions than conventional bench-scale reactors (Amanullah et al., 2001; Enfors et al., 2001; Lara et al., 2009; Onyeaka et al., 2003). Often, SDRs consist of a stirred tank reactor (STR) connected to a plug flow reactor (PFR), in which the cultivation medium is circulated between the two different compartments. Different scenarios of oxygen, pH, and substrate gradients often encountered in large scale fermentation have been simulated in SDRs by selecting appropriate points for oxygen inlet, nitrogen sparging and/or substrate feed. For example, by having the oxygen inlet in the STR and the glucose feed zone in the beginning of the PFR, a glucose rich zone in combination with an oxygen limited zone in the PFR was accomplished. This experimental setup corresponded well to the situation cells experienced in large scale fed-batch *E. coli* cultivation (Enfors et al., 2001). Thus, SDRs become a useful tool to study the response of bacterial cultivations to spatially structured environments.

4.3. Pulse experiments

Another way to study physiological responses to concentration heterogeneities is to perform continuous cultivations and perturb the system with pulses of substrates (Kacmar et al., 2006; Sweere et al., 1988a; Theobald et al., 1997) or with sudden changes in oxygen concentration (Abel et al., 1994; Sweere et al., 1988b). The metabolic and physiological responses to the perturbations together with gene expression analysis can give valuable information for the improvement of the production process through new strain design. Furthermore, implementation of single-cell analysis techniques, such as flow cytometry or molecular techniques (cf. Section 3), can provide insight into how different environmental factors influence cell population heterogeneity and whether there exists such a thing as an optimal level of heterogeneity.

4.4. Microbioreactors

A new experimental tool has become available through the development of microbioreactors (MBR), which typically operate at volumes <1 mL (Betts and Baganz, 2006; Schäpper et al., 2009). The small size gives MBRs some inherent characteristics that are favorable for investigative studies: The gradients encountered in MBRs can be kept quite small as the whole volume can be mixed rather efficiently, which facilitates control of the culture parameters. Mixing can for example be achieved using scale-down versions of conventional impellers, or by novel approaches, such as peristaltic mixing tubes in the ceiling of the reactor chamber of the MBR. Additionally, the small thermal mass of MBRs allows for quick temperature changes of the entire broth. Typically, MBRs are aerated through a membrane, so that no bubbles are present in the system. This greatly increases the signal quality of optical on-line measurement systems, which in turn increases the amount of information that can be gained per experiment. The signal qualities obtained in MBRs are sensitive enough to allow the monitoring of fluorescence/luminescence from reporter strains. In addition to monitoring physicochemical variables like dissolved oxygen and pH, it is therefore possible to monitor in real time the "stress response" systems of a cell. These characteristics together make it possible to simulate the varying conditions a cell might encounter during its journey through a large scale reactor with a high degree of control. This could for example be achieved by using an array of individually controlled MBRs in batch operation mode, or by serially moving through different culture conditions in a continuous culture MBR (micro-chemostat). Due to the availability of on-line measurement tools, the influence of the changing conditions on the cells can then easily be quantified. Finally, the operation volumes of MBRs are still large enough to combine the described non-invasive optical methods with conventional endpoint analyses. Boccazzi et al. (2006), for example, demonstrated that such MBR can be combined with DNA microarrays to obtain differential gene

expression profiles. With the rapid development in microfluidic approaches for cell analysis at the genome, transcriptome, proteome, and metabolome level (Szita et al., 2010), it seems feasible in the future that MBRs can be linked with such analytical methods in an automated fashion. This will enable the rapid investigations of environmental factors responsible for heterogeneity in microbial populations and provide automated and information-rich experiments that underpin novel modeling frameworks (cf. Section 5).

In order to use MBRs as a tool to study phenomena taking place in large-scale reactors, it is essential to assess their scalability. Indeed, the degree of similarity that can be achieved when mimicking varying process conditions from large-scale reactors in microbioreactors depends on which parameters are identified as critical (e.g. gassed power per unit volume, agitator tip speed, constant DOT, oxygen mass transfer capacity-k_La, mixing time), and how well they are controlled in MBRs and how scale-sensitive the process is to begin with (DePalma, 2010). The type of microbioreactor and scale-down parameter to be used for a particular bioprocess should be chosen based on examining cell characteristics and process conditions (Betts and Baganz, 2006; Islam et al., 2008). With regard to, for example, mixing, it is important to obtain in-depth knowledge on basic fluid mechanic for each MBR design. Experimental studies based on microparticle image velocimetry (μPIV) measurements, in combination with different simulation tools (e.g. computational fluid dynamics, cf. Section 5.3.1) can be used to characterize mixing in terms of power input and shear stress in the MBR (Micheletti and Lye, 2006).

4.5. Sampling strategies and bioprocess control

A prerequisite for this kind of scale down experiments is the knowledge of *how* conditions vary in large scale reactors. To be able to have a better understanding of the interaction between biological systems and bioprocess operation conditions, it is essential to have good monitoring of key process variables and consequently a proper sampling procedure. Therefore it is essential to devise a sampling system which can obtain representative samples from the bulk material in large scale, by choosing the proper sampling locations, the number of samples and their size (Holm-Nielsen et al., 2006; Mortensen and Bro, 2006). The ideal sampling procedure should be rapid, on-line, non-destructive, and stable over time, provide easy handling, have a low contamination risk and should not disturb the system (Ritzka et al., 1997). Usually, real-time monitoring in commercial bioreactors is possible for dissolved oxygen, pH, temperature and pressure, but for measuring cell mass, viability, substrate and product concentrations, off-line measurements are still dominating. However, with development of analytical methods and new sensor technology there are possibilities for on-line measurement of different bioprocess variables like substrate and product concentration without withdrawal of samples (in-situ). Optical sensors together with spectroscopic methods offer the advantages of noninvasive, nondestructive, continuous, and simultaneous multi-analyte monitoring (Ulber et al., 2003). Nonetheless, the use of on-line near infrared spectroscopy for monitoring of fermentations is not straightforward, as reviewed by Cervera et al. (2009).

With the development of wireless network technology and the application of inexpensive wireless devices-sensors (motes), which may float along in the cultivation broth, there is a great potential for deeper understanding of the different conditions encountered by the cells. The sampling strategy here is concerned with the number and location of motes (Farré et al., 2009; Nasipuri et al., 2006).

Still, it is important to bear in mind that with application of new biosensors and analytical methods some practical implementation problems could occur (Cervera et al., 2009; Schäpper et al., 2009). In general, the sampling strategy should be adjusted to the monitoring strategy and should be carried out according to the theory of sampling in order to obtain reliable and reproducible experimental data (Gy, 1998; Mattiasson and Håkanson, 1993).

5. Modeling heterogeneous microbial populations

Similar to chemical systems, the design, control and optimization of bioreactors in many academic and industrial applications have been based on macroscopic first principle models built on balances of extensive properties. However, modeling of bioreactors presents additional challenges as a result of the microorganism intrinsic metabolic regulation (Ramkrishna, 2003). Cell variability results in for example non-linearities associated with cell growth and division processes even when operated at constant temperature (Daoutidis and Henson, 2002). Further challenges to the formulation of models with an appropriate level of detail and predictive ability arise thus when dealing with biological systems.

With the continuous development of the capabilities of techniques available for monitoring of cell properties at single-cell level (cf. Section 3) it is nowadays possible to monitor the distribution of cell properties during a microbial cultivation. In order to take advantage of this knowledge and improve the design and control of bioprocesses, mathematical models able to describe the behavior of a dynamic microbial population are necessary in order to support and facilitate the interpretation of the resulting data sets. Furthermore, the validation of such models is made possible by experimentally simulating large scale conditions in smaller set ups as discussed in Section 4.

In this section we discuss models suitable for describing populations of individual microbial cells both in the case of perfectly mixed bioreactors as well as in the case of large scale reactors where gradients are formed due to mixing limitations (cf. Section 4).

5.1. Structured and segregated models

Aiming at describing and predicting the behavior of cell populations, several mechanistic models of various degrees of complexity have been proposed during the last decades. In this section, we start by presenting a classification of such models of cell populations, under the assumption of a homogeneous (i.e. perfectly mixed) reactor environment. A schematic summary is presented in Fig. 6. If the assumption of a homogeneous reactor environment does not hold, then a distributed model, where spatial coordinates form an independent variable in addition to time, is required (Gernaey et al., 2010). The case of spatial heterogeneity will be discussed in Section 5.3.

A very large share of the models proposed for microbial populations is *unsegregated* (Nielsen and Villadsen, 1992), i.e. based on an average cell description. *Unsegregated unstructured* models are the simplest ones as biomass is considered as a black box: Intracellular kinetics are not described, and only the input (e.g. substrate feeding) and output (e.g. substance of interest produced by the microorganism) are accounted for.

Unsegregated structured models form an important class, and incorporate information on the internal mechanism and composition of the microbial mass with the use of several variables, e.g. NADH, precursors, metabolites, ATP, biomass (Gernaey et al., 2010). The number of variables used in such a model should, however, be restricted to a minimum: only variables necessary to obtain information about the most relevant processes of interest should be included. Unsegregated structured models have been used for modeling complex processes, such as yeast intracellular metabolism (e.g. Nielsen and Villadsen, 1992), and morphology-specific growth of filamentous fungi (e.g. Agger et al., 1998). Cybernetic models (e.g. Jones and Kompala, 1999; Young et al., 2004) and different kinds of models based on genomic data (e.g. Becker et al., 2007) belong to this category. These models are beyond the scope of this contribution, and

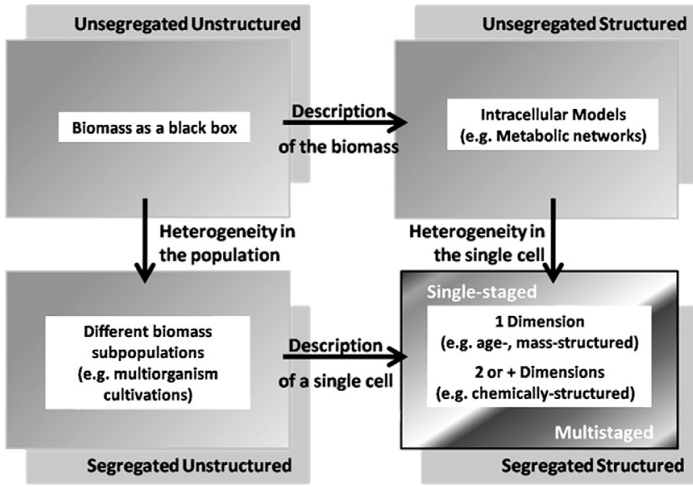

Fig. 6. Schematic classification of mechanistic models for cell cultivations. *Segregated* means cell heterogeneity is taken into account, and *structured* means different cellular components are described.

have been reviewed elsewhere (Gombert and Nielsen, 2000; de Jong, 2002; Stelling, 2004; Teusink and Smid, 2006).

Segregated models account for cell-to-cell variation by considering distributed rather than uniform cell properties (Fredrickson et al., 1970; Fredrickson, 2003). They are, thus, necessarily statistical and the degree of structure is at least a scalar such as cell size (mass or volume) or cell age (Ramkrishna, 2000). A more restrictive classification for *segregated structured* models has been used (Bailey, 1998), where structured refers exclusively to a description of a single cell using multiple biochemical substances (i.e. chemically structured models). In this contribution, we will follow the broader definition proposed by Ramkrishna (2000), who considers that age and cell mass are descriptors of the cell state, even though they may only indirectly be indicators of the cell metabolism. Therefore, by *segregated structured* models (see Fig. 6) we refer to both single variable (e.g. age, mass) models, as well as multivariate, chemically structured models. Consequently, we classify as *segregated unstructured* models the cases where the model describes the behavior and size (i.e. total number of cells) of co-existing subpopulations by considering cells by their existence without any further description of the details of cellular metabolism. This would be, for example, the case of activated sludge models that describe the dynamics of heterotrophic and autotrophic biomass subpopulations (e.g. Gernaey et al., 2004; Henze et al., 2000). In single organism cultivations, subpopulations would, for example, reflect different cell cycle phases. In this case, transitions of cells from one subpopulation to another are possible, and the use of a cell descriptor variable is, therefore, necessary to account for them. This makes it virtually impossible to predict the dynamics of a single organism cultivation using a segregated unstructured model, and further discussion will, thus, focus on segregated structured models.

Segregated structured models can be further classified as single or multi-staged models. The latter account for different cell stages where significant differences in the metabolism are observed, such as budding and non-budding cells (e.g. Cazzador and Mariani, 1988), daughter and parent cells of different generations (e.g. Hatzis and Porro, 2006), different cell cycle phases (e.g. Fredrickson, 2003), productive or non-productive phases (e.g. Mantzaris et al., 2002).

Different formulations are possible for segregated structured models. Population Balance Models (PBMs) provide the most generic approach to modeling distributed properties, and we focus on this type of models in this review. Other simpler models based on ordinary differential equations (ODE) and delay differential equations (DDE) have been reviewed elsewhere (Bley, 2010).

5.2. Population Balance Models for microbial populations

In general, a PBM predicts the temporal change of the cell number distribution, which is characterized by a descriptor variable x (e.g. cell age, mass, intracellular metabolites). This change can result from cell growth and division into newborn cells. Different formulations are used depending on the cell descriptor variable used.

In Tables 2–4 we have briefly described models for microbial populations that have been published in the last four decades. PBMs typically consist of a Population Balance Equation (PBE), along with boundary and initial conditions as well as other coupled equations describing cell division probability and intensity, partitioning of cell content upon division, stage transitions and, in the case of chemically structured models, cellular kinetics. PBEs can be defined as equations of change, i.e. balance equations that account for the various processes that change the number of cells in a population (Fredrickson, 2003), and take the form of first-order partial integro-differential equations, while the supplementary equations, coupled in a non-linear way, are typically ordinary integro-differential equations (Mantzaris et al., 2001a).

To avoid the challenges in solving complex chemically structured PBMs, Mantzaris et al. (2001a) proposed using a large, but finite number of single cells to represent an entire microbial population. In this Monte Carlo approach, a population of single cells, or cell ensemble, is generated by randomizing kinetic parameters or initial conditions for a single cell model, and it is assumed that the continuity of solutions implicit in the PBM solutions, can be simulated if a large enough number of cells are used. Cell ensemble models have been used to describe respiratory and glycolytic oscillations in yeast populations (Henson, 2003, 2004, 2005; Henson et al., 2002), and further discussions on this approach can be found elsewhere (Henson, 2003; Stamatakis, 2010).

5.2.1. Single variable PBMs: formulation of mass- and age-structured models

Two different formulations are typically used for a single variable PBM (i.e. one-dimensional or 1-D PBM): Age- or mass-structured

formulation. The first forms the most simple PBM for microbial cultures as the variation of age with time is unity, and it is, therefore, possible to avoid modeling individual cell growth kinetics (Subramanian et al., 1970). Considering a continuous cultivation in a homogeneous environment, and regarding age as the time elapsed since the birth of the cell, the age-structured PBM expresses the temporal change of the number of cells having a given age. It can be mathematically described by Eqs. (1)–(3), where $N(a,t)da$ is the number of cells with age a at time t, $\Gamma(a)$ is the division rate function and D is the dilution rate of the bioreactor (D will be zero for batch cultivations and positive for continuous and fed-batch operations). The boundary condition (Eq. 2) defines the number of cells with age zero (i.e. newborn cells). Assuming two newborn cells are originated upon division, the boundary condition thus equals twice the total number of cells dividing at a given time instant. Moreover, the total number of cells dividing is given by the integral of the division term in Eq. (1) over the age span. Finally, the initial distribution N_0 is considered to be known (Eq. 3).

$$\frac{\partial N(a,t)}{\partial t} + \underbrace{\frac{\partial N(a,t)}{\partial a}}_{\text{Cells becoming older than a}} = \underbrace{-\Gamma(a)N(a,t)}_{\text{Cells dividing into daughter cells}} - \underbrace{DN(a,t)}_{\text{Cells flowing out of the reactor}} \quad (1)$$

$$N(0,t) = 2\int_0^\infty \Gamma(a)N(a,t)da \quad (2)$$

$$N(a,0) = N_0(a) \quad (3)$$

Although the model is mathematically solvable (e.g. Hjortsø, 1995; Liou et al., 1997), the adequacy of using age as indicator of the organism state has been questioned (Fredrickson, 2003). Indeed, a fundamental question arises: is it feasible to monitor the distribution of cell ages, when investigating the development of heterogeneous populations during a dynamic cultivation?

Cell proliferation studies, commonly used in immunology (Lyons, 2000), have been performed in order to assess cell growth, including lag times, cell division and injury, after treatments with heat and antimicrobial agents (Ueckert et al., 1997). In this type of studies, samples taken from the primary cultivation are stained using fluorescent dye (carboxyfluorescein succinimidyl ester, CFSE), and re-suspended in growth media (secondary cultivations). Based on the fact that the fluorescence intensity decreases by approximately a factor of 2, upon cell division, cell division rates can be estimated based on analysis of fluorescence histograms (obtained by flow cytometry) at different times of the secondary cultivation (Luzyanina et al., 2007). In the context of this review, experimental validation of an age-structured model requires the experimental determination of distributions of cell ages at different time instants of a dynamic cultivation. To our knowledge, it is not obvious that the information obtained by cell proliferation studies, where information is collected during secondary cultivations, can be correlated to age distributions of microbial populations in the primary cultivation.

In fact, the problematic around monitoring of cell age revolves also around the concept of cell age itself. On the one hand, in the case of budding microorganisms, the co-existence of a generation zero of newborn cells, and several generations of mother cells (i.e. of cells which have created one or more daughter cells by budding) could be monitored based on the existence of bud scars and the fact that the cell wall of the newborn cell is synthesized upon budding (e.g. Alberghina et al., 1998). On the other hand, in the case of microorganisms dividing by fission, it is not possible to clearly distinguish mother and daughter cells. The definition of age is thus intimately connected to the progression through the cell cycle. Extensive work on monitoring cell cycle progression using both microscopy and flow cytometry has been published and reviewed elsewhere (Porro et al., 2009).

The use of mass as descriptor variable circumvents, at least partly, these problems as distributions of masses can be easily obtained experimentally by using e.g. flow cytometry (see Section 3.4). Moreover, it has been observed that the size distribution of a cell population does respond to changes in the extracellular environment (Wheals, 1982). In this case *mass* can represent the total cell mass or volume, as well as any conserved property of the cell such as the mass of intracellular components (e.g. total protein, DNA or RNA content). For mass-structured PBMs, the boundary condition (Eq. 5) reflects that there are no cells with mass zero, and, consequently, a birth term is included in the PBE. Additionally, the variation of cell mass with time (i.e. growth rate, r_m) is included in the PBE, and its dependence on the extracellular environment (e.g. available substrate) can be incorporated here (Mantzaris et al., 1999). A mathematical formulation for a mass-structured PBE for a microbial population in a well mixed reactor, along with initial and boundary conditions is provided in Eqs. (4)–(6), where $N(m,t)dm$ is the number of cells with mass m, $\Gamma(m,S)$ is the division rate function, $P(m,m',S)$ is the partitioning probability density function, describing the probability of a mother cell of mass m' to form a daughter cell with a lower mass m upon division, and the initial distribution, $N_0(m)$ is known (Eq. 6).

$$\frac{\partial N(m,t)}{\partial t} + \underbrace{\frac{\partial}{\partial m}[r_m(m,S)N(m,t)]}_{\text{Cells that have increased mass}} = \underbrace{-\Gamma(m,S)N(m,t)}_{\text{Cells dividing into daughter cells}} + \underbrace{2\int_m^\infty \Gamma(m',S)P(m,m',S)N(m',t)dm'}_{\text{Cells being born with mass m}} \underbrace{-DN(a,t)}_{\text{Cells flowing out of the reactor}} \quad (4)$$

$$N(0,t) = 0 \quad (5)$$

$$N(m,0) = N_0(m) \quad (6)$$

Although mass distributions are more easily measurable, mass-structured PBMs are not able to predict common cellular behaviors such as time lags in the response of cells to an extracellular stimulus (Fredrickson, 2003), and this may explain the sparse use of these models for control, design and optimization of bioprocesses (Mantzaris et al., 2001a). One solution to this is the use of more than one descriptor variable to better account for cellular metabolism. Multi-stage models can be regarded as an attempt to take more cell descriptors into consideration without increasing the numerical difficulties associated with a two dimensional PBM. Hatzis et al. (2006) proposed a model where different generations (age stages) are considered, as well as (un)budded sub-stages, and a mass-structured PBM is used to obtain the number distribution for each of the subpopulations. This allowed taking into account the effect of aging on the growth, budding and division while avoiding the complexity inherent to a two-dimensional (2D) PBM. The definition of transition functions for cell division/birth and budding presents nonetheless an increased effort relatively to single-staged models. Experimental validation of this model is provided in Cipollina et al. (2007). A 2D PBM where a continuum approach is used for both age and mass, has been formulated and numerically solved by Liou et al. (1997). Experimental validation was however not performed. The collection of experimental data for a continuous span of age and mass is, in fact, virtually impossible.

5.2.2. Multivariable PBMs: formulation of chemically structured models

Fredrickson et al. (1967) proposed the use of a vectorial description of the cell physiological state, which they designated as a physiological state vector. This descriptor vector consisted of several masses of biochemical substances found in a cell. The mathematical

Table 2
One-dimension Population Balance Models (PBMs).

1-D PBM				
Variable: Age				
Stages	Reactor Mode	Description	Experimental data	Reference
Single	Continuous	PBM where the environmental conditions are the key cell cycle parameter. Periodic oscillations are sustained by periodic change in the environment, without using specific kinetic expressions. Two models are proposed for binary fission organisms and budding yeast.	Predicted oscillation periods were compared with experimentally observed values, for a range of dilution rates.	(Hjortsø and Nielsen, 1995)
Single	Continuous	PBM was used for predicting periodic behavior of a *S. cerevisiae* cell population and its relation to cell-cycle synchrony.	Experimental observations are taken into account in the formulation of the model.	(Zamamiri et al., 2002)
Single	Continuous	PBM for synchronous growth of *S. cerevisiae* with asymmetric budding cycle. The model describes sustained oscillations with constant cell number distributions.	Model validation by comparison with experimental data is presented in a subsequent publication. (Bellgardt, 1994b)	(Bellgardt, 1994a)
Single	Continuous	PBM was used in the design of a controller. Nonlinear feedback control laws are derived in order to attenuate undesired oscillations, or induce synchrony in the *S. cerevisae* culture.	–	(Kurtz et al., 1998)
Active Inactive Dead	Continuous	Model for bioprocess catalyzed by *S. cerevisiae* in a stirred-tank, which is able to reproduce periodic behavior.	Parameter estimation was done using 6 data sets from batch aerated cultivation. Data had been published previously (Duarte et al., 2003).	(Duarte et al., 2003)
Cell cycle phases Labeled/unlabeled	Continuous	PBM for human leukemia cells (Jurkat) which models the two subpopulations generated by addition of bromo deoxyuridine.	Age-dependent model parameters were extracted from portioned population data.	(Sherer et al., 2008)
7 development stages from sporangium to mature cell	Batch Fed-batch	PBM was applied to describe the maturity of sporangium of *Bacillus subtilis* toward the formation of spores. It describes the differentiation phenomenon with associated product formation.	Parameters in the model were determined by fitting the model to experimental data.	(Huang et al., 2003)
1D PBM				
Variable: Mass				
Stages	Reactor mode	Description	Experimental data	Reference
Single	Fed-batch	Compare three simple population models for dynamic simulation of step responses in fed-batch cultures — a simple timer model, a discrete age distribution model which results in a set of ordinary differential equations, and a similar one with discrete mass distribution.		(Takamatsu et al., 1985)
Single	Sequential batches	In this work, a segregated, structured microbial population balance model is formulated and used to numerically simulate the self-cycling fermentation (SCF) process.	Experimental observations taken into account in the formulation of the model. The model outputs were compared and validated with previously published experimental data.	(Godin et al., 1999)
Single	Batch	A numerical solution of the mass structured cell PBM in an environment of changing substrate concentration is presented. It can be applied for any type of single-cell growth rate expression, equal or unequal cell partitioning at cell division, and constant or changing substrate concentration.	Experimental observations are taken into account in the formulation of the model.	(Mantzaris et al., 1999)
Single	Batch startup + continuous	PBM consists of a simple structured description of the extracellular environment, as well it accounts for the three most important metabolic pathways involved in cell growth with glucose substrate of *S. cerevisiae*.	The parameter values were adjusted to achieve qualitative agreement with experimental observations.	(Mhaskar et al., 2002)
Single	Batch Continuous	The model predicts several situations of batch and continuous growth in which the population density and biomass concentration show opposing trends due to significant variation in the cell mass distribution with time.	–	(Subramanian et al., 1970)
Single	Continuous	A controller is designed to stabilize steady-state and periodic solutions by regulating the discretized cell number distribution and the substrate concentration. It is based on a dynamic model for the continuous *S. cerevisiae* cultivation.	–	(Zhu et al., 2000)
Daughter Parent Daughter Parent	Continuous	Model aims at simulating the effect of dilution rate on the mode of oscillation in continuous cultures of asymmetric budding yeast *S. cerevisiae*.	The growth properties of the yeast were analyzed for continuous cultivations. The distribution of parent and daughter cells in the population was determined microscopically after staining the	(Beuse et al., 1998)

Table 2 (*continued*)

1D PBM				
Variable: Age				
Stages	Reactor Mode	Description	Experimental data	Reference
Budded Daughter Parent Budded Unbudded	Continuous	PBM describes the growth of S. cerevisiae in spontaneously synchronized continuous cultures.	bud scars and DNA. The structure of the population was identified using oscillating continuous cultures where the division of the cells is synchronized and detectable by large variation of the on-line measurements (gas exchange rate or heat production rate).	(Duboc and von Stockar, 2000)
Daughter Parents Budded Unbudded	Batch	The model framework couples a morphologically-structured representation of the population with population balance theory to formulate a dynamic model for the size distribution of growing yeast populations.	Model validation by comparison with experimental data is presented in a subsequent publication. (Cippolina et al., 2007)	(Hatzis and Porro, 2006)
Daughter (Small, large) Budded	Continuous	The model established a dynamic PBM for asymmetrically dividing yeast. Three special cases are described: step change in growth rate, two transient behaviors following perturbations in the age-distribution.	It is shown how experimental data on transient behavior of a cell population can yield information on single-cell mass-synthesis kinetics and on the manner in which individual cells control certain critical parameters in the cell cycle.	(Hjortsø and Bailey, 1983)
Budded Unbudded	Continuous	The model describes the structural heterogeneity of yeast cell populations (*S. cerevisiae*) and considers the interaction of the population with its environment. Two different situations were investigated: pulse changes of the dilution rate in a continuous process and of the substrate concentration.	–	(Cazzador and Mariani, 1988)
Budded Unbudded	Continuous	The model aims at understanding the properties of the microbial biomass in terms of its composition and of the regulation of cell growth and division.	The size at bud emergence and the percentage of budded cells was experimentally determined for a range of dilution rates.	(Cazzador et al., 1990)
Cell cycle phases	Batch	A multi-stage population balance model for the growth of ciliated protozoa through its three cell-cycle phases.	Experimental observations are taken into account in the formulation of the model.	(Hatzis et al., 1995)
Non-producing Producing	Continuous	PBM describes the dynamics of cell growth of *S. cerevisiae* during each of the two stages of the cell cycle, including cultivations at limiting substrate and product concentrations.	Experimental observations are taken into account in the formulation of the model.	(Mantzaris et al., 2002)
Single	Fed-batch	See description above.		(Takamatsu et al., 1985)

formulation of such a model is in essence similar to the formulation of a mass-balance (Eqs. 4–6), where m, and m' are now vectors rather than scalars. The implementation of such a highly structured PBM has however never been achieved, due to the complexity of defining the kernel function for growth, division and partitioning upon cell birth, as well as computational tractability issues that arise when attempting to numerically solve such a model.

5.2.3. Growth, division and partitioning functions: The core of the PBM

The difficulties in formulating and solving a PBM lie foremost in the definition of the single cell growth and division (or any stage-to-stage transition) rates, and the corresponding partitioning function. The experimental determination of these intrinsic physiological state functions poses many challenges, mainly as it requires single-cell measurements (Mantzaris et al., 2001a). The experimental observations then have to be translated into mathematical descriptions, very often as statistical distributions.

In the case of mass-structured models without substrate dependency, the mass growth rate for a single cell is typically considered having zero or first order kinetics (Mantzaris et al., 2001a), i.e. $r_m(m) = k_1$ or $r_m(m) = k_2\, m$.

The division rate function is generally described mathematically with the help of a hazard function, $\gamma(m)$ (see Eq. 7) based on the probability density function, $f(m)$, and describing the probability of a cell of mass m to divide in the next time step, given that it has not divided before.

$$\Gamma(m,S) = \gamma(m) r_m(m,S) = \frac{f(m)}{1 - \int_0^m f(m')dm'} r_m(m,S) \tag{7}$$

Although significant progress has been made in the area of single-cell analysis, as discussed in Section 3, it is very often necessary to introduce several assumptions in a theoretical model formulation, which ideally should be validated with experimental data. $f(m)$ is commonly assumed to be a normal distribution with mean corresponding to the critical mass for division (e.g. Mantzaris et al., 1999), while others have used a Weibull distribution (Hatzis and Porro, 2006; Hatzis et al., 1995).

As mentioned in Section 5.2.1, $P(m, m', S)$, describes the probability of a mother cell of mass m' to form a daughter cell with a lower mass, m upon division. For the sake of simplicity, this conditioned probability density function is, typically, assumed to be independent of the available substrate, and has been described as a statistical beta function (e.g. Mantzaris et al., 1999; Hatzis et al., 1995) or a normal distribution (Hatzis and Porro, 2006).

Table 3
Two-dimension Population Balance Models (PBMs).

2-D PBM				
Variable: Age and mass				
Stages	Reactor mode	Description	Experimental data	Reference
Single	Batch	The model describes the production of ethanol in glucose fermentation of *Zimomonas mobilis*.	Model validation was presented in a second publication (Fiolitakis, 1987b).	(Fiolitakis, 1987)
Single	Batch Continuous	The growth-controlled mathematical model of budding yeast predicts theoretical protein and volume distributions.	Compare with protein and volume distributions measured by flow cytometry, for populations growing both in batch and in glucose-limited chemostat cultures.	(Mariani et al., 1986)
Single	Batch	A age and mass structured PBM based on the assumption that only cells from a kth generation originate the (k + 1)th generation. This successive generation approach is applied first to one-dimensional model, and then to 2-D one.	Experimental observations are taken into account in the formulation of the model.	(Liou et al., 1997)

5.2.4. Numerical schemes for solving PBMs

The complexity of the numerical schemes necessary to solve a PBM forms a second explanation for the sparse use of PBMs in optimization or design of bioprocesses (Hatzis and Porro, 2006).

Analytical solutions have been derived for balanced growth cultivations in very specific operational conditions (Bellgardt, 1994a; Hjortsø, 1995; Subramanian et al., 1970). Numerical schemes are, however, necessary to obtain dynamical solutions. Case-specific numerical schemes have been developed such as the three-step procedure (Cazzador and Mariani, 1988) and the successive generation approach (Liou et al., 1997). Other numerical methods generally used for partial differential equations have also been applied, such as finite differences (e.g. Kurtz et al., 1998; Mantzaris et al., 1999, see Fig. 7), finite elements (e.g. Godin et al., 1999; Kavousanakis et al., 2009), orthogonal collocation (Mhaskar et al., 2002; Zhu et al., 2000) and spectral methods (e.g. Mantzaris and Daoutidis, 2004). Furthermore, Hatzis et al. (1995) proposed a Monte Carlo procedure for solving a multi-staged mass-structured PBM, and Huang et al. (2003) solved a multi-staged age-structured model using the method of characteristics. A detailed discussion of the mentioned numerical methods is beyond the scope of this contribution. Mantzaris et al. (2001a,b,c) provide a comprehensive review and analysis of finite differences schemes, spectral methods, and finite element schemes for solving multivariate PBMs.

5.3. Modeling spatial heterogeneity

The physiological state of cellular systems and its impact on growth and product formation is the result of a complex interplay between the extracellular environment and the intracellular machinery (Mantzaris et al., 2001a,b,c). As discussed earlier, cells are subjected to spatio-temporal variations in large scale reactors, unlike in

Table 4
Multi-dimensional Population Balance Models (PBMs).

Multi-dimensional PBM				
Variable: Physiological state vector				
Stages	Reactor mode	Description	Experimental data	Reference
Single	Batch Continuous	First formulation of a multidimensional PBM: introduction of the concept of physiological state vector.	–	(Fredrickson et al., 1967)
Single	Continuous	The model aimed at studying the existence of self-similar forms (e.g. time invariant) when each physiological state is scaled with respect to its population average. In this article, each physiological entity was scaled with the respective population average of that entity.	Experimental observations are taken into account in the formulation of the model	(Ramkrishna, 1994)
Single	Continuous	A controller is formulated having the PBM as base model. It aims at controlling different moments of the cell mass distribution in a continuous bioreactor by manipulating the dilution rate.	The use of flow cytometry combined with available staining techniques, which allow the on-line measurement of cell property distributions can make the practical implementation of such a control approach possible.	(Mantzaris and Daoutidis, 2004)
Multi-staged	Batch	A new and different approach involving randomization of growth rates and compartmentalization is proposed. It aims at circumventing the necessity of having intensity functions for transitions between cell cycle phases, and for which the fission intensity function is state-independent.	–	(Fredrickson and Mantzaris, 2002)
Cell cycle phases	Batch	The model is a generalization of the first multidimensional PBM (Fredrickson et al., 1967), which accounts for passages of cells through a series of recognizable cell cycle phases.	–	(Fredrickson, 2003)

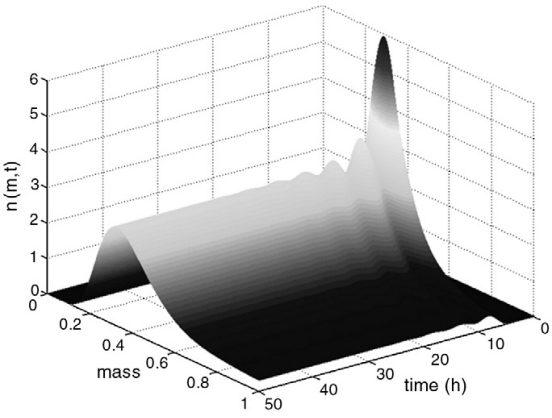

Fig. 7. Normalized cell mass distribution as a function of time obtained for a 1-D mass structured PBM. A hybrid numerical scheme was used to solve the PBM (based on Mantzaris et al., 1999). The mass units do not have a physical meaning.

laboratory-scale studies under controlled conditions. Indeed, when an individual microorganism circulates through a large scale reactor, it is sequentially exposed to these different local conditions (Pollard et al., 1998). This may significantly influence the behavior of cellular processes and make conventional (i.e. which assume homogeneous environments) models inapplicable (Larsson et al., 1996).

Coupling fluid models for the extracellular environment with models for the cell population allows for reflecting the interaction between the environment and the physiological state of the cell (Larsson et al., 1996). A framework suited for capturing the local and global variations in both intra- and extracellular concentrations relies on the link between metabolic network modeling and computational fluid dynamics (CFD) (Lapin et al., 2004). The use of CFD in modeling bioprocesses is thus gaining importance, both in academia and industry (Bezzo et al., 2005; Fang, 2010a,b; Generalis and Glover, 2005).

5.3.1. Integration of computational fluid dynamics (CFD)

Computational fluid dynamics has proven to be an efficient and powerful tool for the design and optimization of several flow applications. Typically, a CFD model consists of the three fundamental equations of fluid flow under a given set of conditions to be solved: continuity, momentum and energy equations. Due to the complexity of these equations they are solved numerically to describe the behavior of the system. It is possible to obtain precise predictions of flow and reaction variables using CFD that can be used in scale-up and design applications. Recently, there has been an increased interest in applications of CFD in the (bio)pharmaceutical and biotechnology industries (Bezzo et al., 2005; Fang, 2010a,b; Generalis and Glover, 2005). This includes analysis of turbulent flow patterns, energy dissipation rates, as well as heat and mass transfer in bioreactors (Zhang et al., 2009), chemical reactions and phase transitions (Lapin et al., 2004) involving several unit operations such as fermentation, mixing and filtration.

Currently, there are two widely used computational approaches for modeling the interaction between phases (Kelly, 2008): the Euler–Euler approach in which different phases are treated mathematically as interpenetrating continua (Barrue et al., 2001; Micale et al., 2000), and the Euler–Lagrange approach in which the fluid phase is treated as a continuum whereas the dispersed phase is solved by tracking a large number of particles through the calculated flow field (Decker and Sommerfeld, 1996; Derksen, 2003). Bezzo et al. (2003) studied xanthan gum production in stirred tanks, and combined the Eulerian approach for the fluid phases with a multizonal model in which the reactor was divided into a limited number of spatial regions. Elqotbi et al. (2006) implemented an Euler–Euler multi-fluid model to study the interaction of fluid flow, mass transfer and reaction in the fermentation of gluconic acid by *A. niger* in a gas–liquid stirred fermenter. They assumed a constant bubble size thus limiting the possibility of predicting local mass transfer across the phases. Using the Eulerian–Lagrangian approach, the interaction between the intracellular state of the individual cells of the population and the turbulent flow fields has been studied in a 68 L (Lapin et al., 2004; see Fig. 8) and a 900 L bioreactor (Lapin et al., 2006). They used both structured segregated and unstructured unsegregated approaches for modeling the biophase, with a 3-D CFD simulation for the reactor. Although Lapin et al. (2004, 2006) were successful in accounting for the interaction between the individual cells and the spatial concentration gradient caused mainly due to the turbulent flow field, using the stochastic Lagrangian approach — a large number of cells ($\sim 10^5$) was required to achieve a realistic description of the population, which is computationally intensive.

As the flow in an agitated bioreactor is usually turbulent, the Reynolds-Averaged Navier–Stokes (RANS) model (Jenne and Reuss, 1999; Lapin et al., 2004) and the Large Eddy Simulation (LES) model (Revstedt et al., 1998) are often used to model the turbulence. The RANS equations govern the transport of the averaged flow quantities, with the whole range of the scales of turbulence being modeled. In LES, large eddies are explicitly resolved in a time-dependent simulation using the 'filtered' Navier–Stokes equations.

A dynamic simulation of a heterogeneous cell population in a non-homogeneous environment can be achieved by coupling CFD and PBM

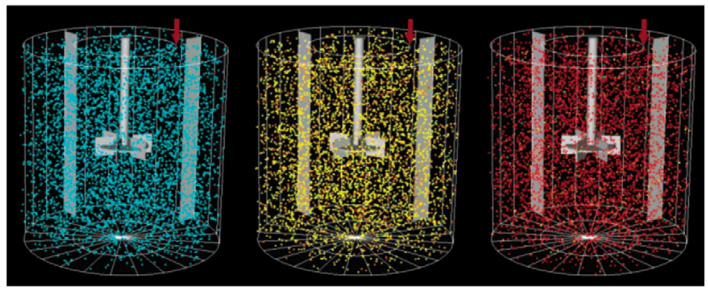

Fig. 8. Movement of a population of 100 000 oscillating yeast cells traveling through the three-dimensional turbulent flow field of a 68-L stirred-tank bioreactor. The prismatic colors of the particles indicate the intracellular NADH concentrations, which ascend from magenta (0.1 mM) via cyan, green, and yellow to red (0.22 mM). (Lapin et al., 2004).

into an integrated framework. The PBM describes the development of the microbial population in a given extracellular environment, while the CFD component allows for the determination of local environmental conditions by calculating flow streams and cell trajectories within the reactor. The integrated CFD–PBM framework should be able to predict local distributed cell properties (e.g. size, composition, age, growth rates, product formation rates) of microorganism populations, while accounting for the changes in the cell physiological state due to different physical (e.g. local gas bubble size, shear stress) and chemical (e.g. local substrate concentration, pH) environments. As discussed previously, the solution of multivariate PBMs requires the use of complex numerical methods. These computational issues are further aggravated with the integration with a CFD model, which would lead to higher calculation times. Nevertheless, it is foreseen that the insight gained by implementing such detailed models will, in the long term, translate into more efficient bioreactor operation (Gernaey et al., 2010).

6. Conclusion and outlook

Although biomass has been typically considered as an abstract entity characterized by average properties in the design, optimization and control of bioprocesses, it is a given fact that even isogenetic (i.e. with the same genetic content) microbial populations have a certain degree of heterogeneity. Indeed, individual microorganisms, even if part of a 'clonal' or isogenetic population, may differ greatly in terms of genetic composition, physiology, biochemistry, or behavior (Brehm-Stecher and Johnson, 2004). This heterogeneity results from differences in the microenvironment surrounding each individual cell, as well as the physiological stage of an individual cell when subjected to a given change in the extracellular medium. Moreover, heterogeneity has been considered to boost the fitness of populations (Avery, 2006), as well as being the cause for the varying degrees of resistance to antimicrobial treatments (Dhar and McKinney, 2007; Sumner and Avery, 2002) and other stresses such as sudden changes in substrate concentration or pH, as well as for the decrease in productivity observed in the scale-up of fermentation processes (Enfors et al., 2001).

To understand the development of heterogeneity, as well as to be able to establish monitoring strategies that can quantitatively reflect the heterogeneity of a microbial population, and, in the future, may yield an increased process knowledge within the bioprocess industry, three essential issues must be addressed: 1) the development of suitable reporter systems and analytical methods that allow for monitoring the changes of the physiological state of an individual cell in a systematic and, ideally, automatic fashion; 2) the study of the cause-effect of changes in the extracellular environment on the single cell physiological state starting at small scale, where it can be assumed that all cells are subjected to the same environmental conditions, and progressively scaling up to large reactors, where single cell trajectories through the reactor should ideally be considered individually due to the existence of gradients; and 3) the formulation and solution of computationally tractable models which are able to account for and predict distributions of cell properties for a heterogeneous cell population rather than being based on average measurements that mask the dynamic structure of the population (cf. Fig. 1).

In this contribution, we attempted to cover the three key topics by presenting and discussing experimental approaches which deliver quantitative information on the distribution of single cell properties in a heterogeneous microbial population, cultivation strategies that allow for understanding of the dynamic environment–cell interaction at different reactor scales, and model frameworks which are able to account for distributed cell properties, and may, in the future, be used to achieve a better control of bioprocesses.

Due to cell individuality, monitoring of a microbial population implies the use of analytical methods that can provide information at single-cell level. In fact, single-cell analytical methods are essential to the understanding of connections between cellular biochemistry and behavior, as well as population-level phenomena (Brehm-Stecher and Johnson, 2004). If a comprehensive study of the sources and mechanisms of cell heterogeneity requires the use of global analytical methods, which for example allow for the evaluation of genetic variability and for understanding the genetic mechanisms cells use to respond to their environment (cf. Section 3.1), these methods are not suitable to monitor a bioprocess. In this case, less time-consuming and more automatic analytical methods are necessary.

Reporter systems (cf. Section 3.2), where only information on relevant selected genes is gathered, consist of a simpler method that can convey information on gene transcription frequency, translation or gene dose, as well as on formation of protein complexes, polypeptide folding, and protein stability. One widely used reporter system relies on the bioluminescence of a GFP which can be used as fusion tags to localize protein and organelles within cells, besides its use as reporter of promoter activity. Furthermore, GFP reporters could be used for monitoring of cell growth in a microbial cultivation at single-cell level by fusing GFP to growth dependent genes, acting as cell stress indicators, as well as pH probes.

Both, fluorescence based microscopy, where the use of fluorescent dyes allows for measuring properties of individual cells such as viability, activity, surface components characterization and internal storage compounds detection, as well as light microscopy, where the use of radiolabeled substrates provides quantitative data on substrate uptake at single cell level, together with the use of image analysis present attractive alternatives for monitoring the physiological state of microorganisms during cultivation (cf. Section 3.3). The same fluorescence stains can be used for flow cytometric analysis (cf. Section 3.4), where data on cell size and granularity distributions can be obtained from light scattering patterns. In addition to automatic data handling providing number distributions of cell properties for a population (cf. Fig. 5), some flow cytometers also offer the possibility of cell sorting (FACS), making flow cytometry a very attractive method for at-line monitoring of microbial populations.

The increasing interest in on-line monitoring of bioprocesses has fostered the development of spectroscopic methods for monitoring of microbial cultivations (Dabros et al., 2008). Raman spectroscopy (cf. Section 3.5) relies on the information provided by molecular vibrations, and is able to deliver a highly specific fingerprint of the molecular structure and biochemical composition of individual cells (Krafft et al., 2009). Therefore, this method may hold promise as an alternative method for single-cell analysis without requiring the use of external markers such as stains or radioactive labels.

As cell heterogeneity is mostly driven by changes in the extracellular environment, it becomes important to understand this interaction, especially in the case of large-scale cultivations. At large scale, concentration gradients are formed due to mixing limitations, and cells are subjected to a series of different microenvironments along their trajectories within the reactor (cf. Section 4). Spatially homogeneous conditions can be studied at micro or lab scale, where the use of microbioreactors (Schäpper et al., 2009) or scale-down reactors has to be considered in order to simulate conditions encountered at large scale.

Metabolic engineering has been used to minimize the effect of environmental gradients on microbial populations, as reviewed by Lara et al. (2006a). For example, the production of fluorescent protein (GFP) by *E. coli* has been improved, by designing metabolically engineered strains which have reduced sensitivity to dissolved oxygen gradients (Lara et al., 2006b). Higher cell viability, also for *E. coli* cultivations, has however been observed in cultivations where glucose, DOT and pH gradients have been simulated (Hewitt et al., 2000; Onyeaka et al., 2003). It is thus our hypothesis that there may be an optimal degree of heterogeneity which corresponds to the ideal trade-off between productivity and viability.

The development of new control strategies, which ensure the cultivation conditions that lead to an optimal degree of cell heterogeneity, requires the existence of models which account for distributed cell properties (cf. Section 5). This type of models is divided into structured and segregated models, and most of the proposed models are population balance models (PBMs). Only very few cell properties can, however, be considered in this type of models without facing numerical challenges and computational tractability issues. In the cases where spatial gradients are to be considered, PBMs can be integrated with computational fluid dynamics (CFD) in order to obtain a model, which simulates the distribution of extracellular medium components as well as the cell trajectories in the reactor. The formulation and validation of such models is not obvious, and requires a comprehensive understanding not only of the mathematics of the models but also of cellular phenomena, and analytical methods used for monitoring microbial populations at single-cell level. However, in our opinion the effort in developing such advanced models is justified. Indeed, such models can be seen as a structured representation of the knowledge available on the interactions between the extracellular environment, the heterogeneity and the productivity of a cultivation. The construction of such advanced models, the confrontation of those models with the available data, and the subsequent use of advanced model analysis tools and optimal experimental design techniques to develop new experiments will, according to us, be important in order to unravel the detailed mechanisms steering heterogeneity.

The amount and quality of the data we are nowadays able to collect at single-cell level is considerable (cf. Section 3), but models for microbial populations, suitable to be used for the optimization and control of bioprocesses, are still virtually inexistent. In fact, there seems to be a gap between the advances of the experimental techniques and the development of models that reflect the observations gathered through these experimental efforts. This gap is most likely resulting from the difficulties in fully understanding the complex cellular mechanisms behind the development of different types of cell heterogeneity.

Moreover, the number of publications resulting from collaborations between experimentalists and modelers is limited. It is, though, in this cooperation, that we believe the key for achieving a higher level of process understanding within the bioprocess industry lies in. Experimental studies are the basis to the understanding of the development of heterogeneous cell populations, but also models, validated using experimental data sets, are necessary to be able to actively use the gathered knowledge in the optimization and control of the bioprocesses.

Acknowledgments

The Danish Council for Strategic Research is gratefully acknowledged for financial support in the frame of the project “Towards robust fermentation processes by targeting population heterogeneity at microscale” (project number 09-065160).

ERA-IB (ERA-NET Industrial Biotechnology) is gratefully acknowledged for financial support in the frame of the project “Targeting population heterogeneity at microscale for robust fermentation processes” (project number EIB.08.031)

References

Abel C, Hübner U, Schügerl K. Transient behaviour of baker's yeast during enforced periodical variation of dissolved oxygen concentration. J Biotechnol 1994;32:45–57.

Adav SS, Lin JC, Yang Z, Whiteley CG, Lee D, Peng X, et al. Stereological assessment of extracellular polymeric substances, exo-enzymes, and specific bacterial strains in bioaggregates using fluorescence experiments. Biotechnol Adv 2010;28:255–80.

Aertsen A, Michiels CW. Stress and how bacteria cope with death and survival. Crit Rev Microbiol 2004;30:263–73.

Agger T, Spohr AB, Carlsen M, Nielsen J. Growth and product formation of *Aspergillus oryzae* during submerged cultivations: verification of a morphologically structured model using fluorescent probes. Biotechnol Bioeng 1998;57:321–9.

Ahn B, Rhee SG, Stadtman ER. Use of fluorescein hydrazide and fluorescein thiosemicarbazide reagents for the fluorometric determination of protein carbonyl groups and for the detection of oxidized protein on polyacrylamide gels. Anal Biochem 1987;161:245–57 1998.

Alberghina L, Smeraldi C, Ranzi BM, Porro D. Control by nutrients of growth and cell cycle progression in budding yeast, analyzed by double-tag flow cytometry. J Bacteriol 1998;180:3864–72.

Alivisatos AP. Semiconductor clusters, nanocrystals, and quantum dots. Science 1996;271:933–7.

Allegra S, Berger F, Berthelot P, Grattard F, Pozzetto B, Riffard S. Use of flow cytometry to monitor *Legionella* viability. Appl Environ Microbiol 2008;74:7813–6.

Amanullah A, McFarlane CM, Emery AN, Nienow AW. Scale-down model to simulate spatial pH variations in large-scale bioreactors. Biotechnol Bioeng 2001;73:390–9.

Arber W, Naas T, Blot M. Generation of genetic diversity by DNA rearrangements in resting bacteria. FEMS Microbiol Ecol 1994;15:5-13.

Atfield PV, Choi HY, Veal DA, Bell PJL. Heterogeneity of stress gene expression and stress resistance among individual cells of *S. cerevisiae*. Mol Microbiol 2001;40:1000–8.

Avery SV. Microbial cell individuality and the underlying sources of heterogeneity. Nat Rev Microbiol 2006;4:577–87.

Bae PK, So H, Kim KN, You HS, Choi KS, Kim CH, et al. Simple route for the detection of *Escherichia coli* using quantum dots. Biochip J 2010;4:129–33.

Bahl M, Sørensen SJ, Hansen LH. Determination of plasmid stability in *Escherichia coli* by flow cytometry. FEMS Microbiol Lett 2004;232:45–9.

Bailey JE. Mathematical modeling and analysis in biochemical engineering: past accomplishments and future opportunities. Biotechnol Prog 1998;14:8-20.

Balaban NQ, Merrin J, Chait R, Kowalik L, Leibler S. Bacterial persistence as a phenotypic switch. Science 2004;305:1622–5.

Barrue H, Bertrand J, Cristol B, Xuereb C. Eulerian simulation of dense solid–liquid suspension in multi-stage stirred vessel. J Chem Eng Jpn 2001;34:585–94.

Bashashati A, Brinkmand RR. A survey of flow cytometry data analysis methods. Adv Bioinform 2009 Article ID 584603, 19 pages.

Beaumont HJE, Gallie J, Kost C, Ferguson GC, Rainey PB. Experimental evolution of bet hedging. Nature 2009;462:90–3.

Becker SA, Feist AM, Mo ML, Hannum G, Palsson BO, Herrgard MJ. Quantitative prediction of cellular metabolism with constraint-based models: the cobra toolbox. Nat Protoc 2007;2:727–38.

Bellgardt K-H. Analysis of synchronous growth of baker's yeast. Part I: development of a theoretical model for sustained oscillations. J Biotechnol 1994a;35:19–33.

Bellgardt K-H. Analysis of synchronous growth of baker's yeast. Part II: Comparison of model prediction and experimental data. J Biotechnol 1994b;35:35–49.

Ben-Amor K, Heilig H, Smidt H, Vaughan EE, Abee T, de Vos WM. Genetic diversity of viable, injured, and dead fecal bacteria assessed by fluorescence-activated cell sorting and 16s rRNA gene analysis. Appl Environ Microbiol 2005;71:4679–89.

Bergquist P, Hardiman E, Ferrari B, Winsley T. Applications of flow cytometry in environmental microbiology and biotechnology. Extremophiles 2009;13:389–401.

Berney M, Hammes F, Bosshard F, Weilenmann H, Egli T. Assessment and interpretation of bacterial viability by using the LIVE/DEAD BacLight kit in combination with flow cytometry. Appl Environ Microbiol 2007;73:3283–90.

Betts JI, Baganz F. Miniature bioreactors: current practices and future opportunities. Microb Cell Fact 2006;5:21.

Beuse M, Bartling R, Kopmann A, Diekmann H, Thoma M. Effect of the dilution rate on the mode of oscillation in continuous cultures of *Saccharomyces cerevisiae*. J Biotechnol 1998;61:15–31.

Bezzo F, Macchietto S, Pantelides CC. General hybrid multizonal/CFD approach for bioreactor modeling. AIChE J 2003;49:2133–48.

Bezzo F, Macchietto S, Pantelides CC. Computational issues in hybrid multizonal/computational fluid dynamics models. AIChE J 2005;51:1169–77.

Bley T. From single cells to microbial population dynamics: modelling in biotechnology based on measurements of individual cells. Adv Biochem Eng 2010:1-17.

Boccazzi P, Zhang Z, Kurosawa K, Szita N, Bhattacharya S, Jensen KF, et al. Differential gene expression profiles and real-time measurements of growth parameters in *Saccharomyces cerevisiae* grown in microliter-scale bioreactors equipped with internal stirring. Biotechnol Prog 2006;22:710–7.

Booth IR. Stress and the single cell: intrapopulation diversity is a mechanism to ensure survival upon exposure to stress. Int J Food Microbiol 2002;78:19–30.

Boulos L, Prévost M, Barbeau B, Coallier J, Desjardins R. LIVE/DEAD® BacLight™: application of a new rapid staining method for direct enumeration of viable and total bacteria in drinking water. J Microbiol Methods 1999;37:77–86.

Brauer MJ, Huttenhower C, Airoldi EM, Rosenstein R, Matese JC, Gresham D, et al. Coordination of growth rate, cell cycle, stress response, and metabolic activity in yeast. Mol Biol Cell 2008;19:352–67.

Brehm-Stecher BF, Johnson EA. Single-cell microbiology: tools, technologies, and applications. Microbiol Mol Biol Rev 2004;68:538–59.

Brock TD, Brock ML. Autoradiography as a tool in microbial ecology. Nature 1966;209:734–6.

Brock ML, Brock TD. The application of micro-autoradiographic techniques to microbial ecology. Theor Ang Limnol 1968;15:1-29.

Bunthof CJ, van den Braak S, Breeuwer P, Rombouts FM, Abee T. Rapid fluorescence assessment of the viability of stressed *Lactococcus lactis*. Appl Environ Microbiol 1999;65:3681–9.

Burmølle M, Hansen L, Oregaard G, Sørensen S. Presence of n-acyl homoserine lactones in soil detected by a whole-cell biosensor and flow cytometry. Microb Ecol 2003;45:226–36.

Burmølle M, Hansen L, Sørensen S. Use of a whole-cell biosensor and flow cytometry to detect AHL production by an indigenous soil community during decomposition of litter. Microb Ecol 2005;50:221–9.
Burmølle M, Hansen LH, Sørensen SJ. Establishment and early succession of a multi-species biofilm composed of soil bacteria. Microb Ecol 2007;54:352–62.
Bylund F, Collet E, Enfors S, Larsson G. Substrate gradient formation in the large-scale bioreactor lowers cell yield and increases by-product formation. Bioprocess Biosyst Eng 1998;18:171–80.
Cabantous S, Pédelacq J, Mark BL, Naranjo C, Terwilliger TC, Waldo GS. Recent advances in GFP folding reporter and split-GFP solubility reporter technologies. Application to improving the folding and solubility of recalcitrant proteins from mycobacterium tuberculosis. J Struct Funct Genomics, 6. ; 2005. p. 113–9.
Cangelosi GA, Brabant WH. Depletion of pre-16s rRNA in starved *Escherichia coli* cells. J Bacteriol 1997;179:4457–63.
Carman KR. Microautoradiographic detection of microbial activity. Handbook of methods in aquatic microbial ecology. CRC Press; 1993. p. 397–403.
Carneiro S, Amaral AL, Veloso ACA, Dias T, Peres AM, Ferreira EC, et al. Assessment of physiological conditions in *E. coli* fermentations by epifluorescent microscopy and image analysis. Biotechnol Prog 2009;25:882–91.
Caro A, Gros O, Got P, De Wit R, Troussellier M. Characterization of the population of the sulfur-oxidizing symbiont of *Codakia orbicularis* (bivalvia, lucinidae) by single-cell analyses. Appl Environ Microbiol 2007;73:2101–9.
Cazzador L, Mariani L. A simulation program based on a structured population model for biotechnological yeast processes. Appl Microbiol Biotechnol 1988;29: 198–202.
Cazzador L, Mariani L, Martegani E, Alberghina L. Structured segregated models and analysis of self-oscillating yeast continuous cultures. Bioprocess Biosyst Eng 1990;5:175–80.
Cervera AE, Petersen N, Lantz AE, Larsen A, Gernaey KV. Application of near-infrared spectroscopy for monitoring and control of cell culture and fermentation. Biotechnol Prog 2009;25:1561–81.
Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC. Green fluorescent protein as a marker for gene expression. Science 1994;263:802–5.
Chalmers NI, Palmer RJ, Du-Thumm L, Sullivan R, Shi W, Kolenbrander PE. Use of quantum dot luminescent probes to achieve single-cell resolution of human oral bacteria in biofilms. Appl Environ Microbiol 2006;73:630–6.
Champe SP, Benzer S. An active cistron fragment. J Mol Biol 1962;4:288–92.
Chien CT, Bartel PL, Sternglanz R, Fields S. The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. PNAS 1991;88: 9578–82.
Choo-Smith L, Maquelin K, van Vreeswijk T, Bruining HA, Puppels GJ, Thi NAN, et al. Investigating microbial (micro)colony heterogeneity by vibrational spectroscopy. Appl Environ Microbiol 2001;67:1461–9.
Christensen BB, Sternberg C, Molin S. Bacterial plasmid conjugation on semi-solid surfaces monitored with the green fluorescent protein (GFP) from *Aequorea victoria* as a marker. Gene 1996;173:59–65.
Chung H, Bang W, Drake M. Stress response of *Escherichia coli*. Compr Rev Food Sci F 2006;5:52–64.
Cipollina C, Vai M, Porro D, Hatzis C. Towards understanding of the complex structure of growing yeast populations. J Biotechnol 2007;128:393–402.
Corich V, Soldati E, Giocomini A. Optimization of fluorescence microscopy techniques for the detection of total and viable lactic acid bacteria in whey starter cultures. Ann Microbiol 2004;54:335–42.
Cormack BP, Valdivia RH, Falkow S. FACS-optimized mutants of the green fluorescent protein (GFP). Gene 1996;173:33–8.
Crameri A, Whitehorn EA, Tate E, Stemmer WP. Improved green fluorescent protein by molecular evolution using DNA shuffling. Nat Biotechnol 1996;14:315–9.
Créach V, Baudoux A, Bertru G, Rouzic BL. Direct estimate of active bacteria: CTC use and limitations. J Microbiol Methods 2003;52:19–28.
Cronin U, Wilkinson M. Monitoring growth phase-related changes in phosphatidylcholine-specific phospholipase C production, adhesion properties and physiology of *Bacillus cereus* vegetative cells. J Ind Microbiol Biotechnol 2008;35:1695–703.
Czechowska K, Johnson DR, van der Meer JR. Use of flow cytometric methods for single-cell analysis in environmental microbiology. Curr Opin Microbiol 2008;11: 205–12.
Dabros M, Schenk J, Marison I, von Stockar U. The ongoing quest for truly on-line bioprocess monitoring using spectroscopy. In: Flynne WG, editor. Biotechnol. Bioeng. New York: Nova Science Publishers, Inc.; 2008. p. 99-119.
Daims H, Wagner M. Quantification of uncultured microorganisms by fluorescence microscopy and digital image analysis. Appl Microbiol Biotechnol 2007;75: 237–48.
Daims H, Lücker S, Wagner M. Daime, a novel image analysis program for microbial ecology and biofilm research. Environ Microbiol 2006;8:200–13.
Daoutidis P, Henson M. Dynamics and control of cell populations in continuous bioreactors. AIChE Symp S 2002;326:393–402.
Davey HM, Davey CL. Multivariate data analysis methods for the interpretation of microbial flow cytometric data. Adv Biochem Eng Biotechnol 2010:1-27.
Davey H, Kell D. Flow cytometry and cell sorting of heterogeneous microbial populations: the importance of single-cell analyses. Microbiol Rev 1996;60: 641–96.
Davey HZ, Winson MK. Using flow cytometry to quantify microbial heterogeneity. Curr Issues Mol Biol 2003;5:9-15.
De Angelis M, Bini L, Pallini V, Cocconcelli PS, Gobbetti M. The acid-stress response in *Lactobacillus sanfranciscensis* CB1. Microbiology 2001;147:1863–73.
de Jong H. Modeling and simulation of genetic regulatory systems: a literature review. J Comput Biol 2002;9:67-103.
Decker S, Sommerfeld M. Calculation of particle suspension in agitated vessels with the Euler–Lagrange approach. Inst Chem E 1996;140:71–82.
Delvigne F, Boxus M, Ingels S, Thonart P. Bioreactor mixing efficiency modulates the activity of a prpos: GFP reporter gene in *E. coli*. Microb Cell Fact 2009;8:15.
DePalma A. Microbioreactors Carve Out Growing Niche. Genet Eng Biotechn N 2010;30:3.
Derksen JJ. Numerical simulation of solids suspension in a stirred tank. AIChE J 2003;49: 2700–14.
Dhar N, McKinney JD. Microbial phenotypic heterogeneity and antibiotic tolerance. Curr Opin Microbiol 2007;10:30–8.
Diáz M, Herrero M, García LA, Quirós C. Application of flow cytometry to industrial microbial bioprocesses. Biochem Eng J 2010;48:385–407.
Duarte M, Medeiros J, Araújo O, Coelho M. An age-structured population balance model for microbial dynamics. Braz J Chem Eng 2003;20:1–6.
Duboc P, von Stockar U. Modeling of oscillating cultivations of *Saccharomyces cerevisiae*: identification of population structure and expansion kinetics based on on-line measurements. Chem Eng Sci 2000;55:149–60.
Elqotbi M, Montastruc L, Vlaev SD, Nikov I. CFD stimulation of gluconic acid production in a stirred gas–liquid fermenter Contribution in: MIXING 2006 12th European Conference on Mixing; 2006.
Enfors SO, Jahic M, Rozkov A, Xu B, Hecker M, Jürgen B, et al. Physiological responses to mixing in large scale bioreactors. J Biotechnol 2001;85:175–85.
Falcioni T, Papa S, Gasol JM. Evaluating the flow-cytometric nucleic acid double-staining protocol in realistic situations of planktonic bacterial death. Appl Environ Microbiol 2008;74:1767–79.
Fang Z. Applying computational fluid dynamics technology in bioprocesses-part 1. Biopharm Int 2010a;23(4):38–43.
Fang Z. Applying computational fluid dynamics technology in bioprocesses-part 2. Biopharm Int 2010b;23(5):42–52.
Farré M, Kantiani L, Pérez S, Barceló D, Barceló D. Sensors and biosensors in support of EU directives. TRAC Trends Anal Chem 2009;28:170–85.
Fearon ER, Finkel T, Gillison ML, Kennedy SP, Casella JF, Tomaselli GF, et al. Karyoplasmic interaction selection strategy: a general strategy to detect protein–protein interactions in mammalian cells. PNAS 1992;89:7958–62.
Fields S, Song O. A novel genetic system to detect protein–protein interactions. Nature 1989;340:245–6.
Finkel SE, Kolter R. Evolution of microbial diversity during prolonged starvation. PNAS 1999;96:4023–7.
Fiolitakis E. Ein altersstrukturiertes populationsmodell zur beschreibung instationärer mikrobieller prozesse, teil I: theorie. Chem Ing Tech 1987a;59:810–1.
Fiolitakis E. Ein altersstrukturiertes Populationsmodell zur Beschreibung instationärer mikrobieller Prozesse, Teil II: Modell-Verifikation am Beispiel der Glucose-Fermentation mit Zymomonas mobilis. Chem Ing Tech 1987b;59:812–3.
Fornasari ME, Rossetti L, Remagni C, Giraffa G. Quantification of *Enterococcus italicus* in traditional Italian cheeses by fluorescence whole-cell hybridization. Syst Appl Microbiol 2008;31:223–30.
Forster S, Snape JR, Lappin-Scott HM, Porter J. Simultaneous fluorescent gram staining and activity assessment of activated sludge bacteria. Appl Environ Microbiol 2002;68:4772–9.
Fredrickson AG. Population balance equations for cell and microbial cultures revisited. AIChE J 2003;49:1050–9.
Fredrickson AG, Mantzaris NV. A new set of population balance equations for microbial and cell cultures. Chem Eng Sci 2002;57:2265–78.
Fredrickson A, Ramkrishna D, Tsuchiya H. Statistics and dynamics of procaryotic cell populations. Math Biosci 1967;1:327–74.
Fredrickson AG, Megeee III RD, Tsuchiya H. Mathematical models in fermentation processes. Adv Appl Microbiol 1970;13:419–65.
Freese HM, Karsten U, Schumann R. Bacterial abundance, activity, and viability in the eutrophic river Warnow, northeast Germany. Microb Ecol 2006;51:117–27.
Fu X, Huang K, Liu S. A rapid and universal bacteria-counting approach using cdse/zns/sio2 composite nanoparticles as fluorescence probe. Anal Bioanal Chem 2009;396: 1397–404.
Gasol JM, Arstegui J. Cytometric evidence reconciling the toxicity and usefulness of CTC as a marker of bacterial activity. Aquat Microb Ecol 2007;46:71–83.
Generalis SC, Glover GMC. Modelling a biochemical reaction with computational fluid dynamics. Int J Chem React Eng 2005;3:A50.
George S, Larsson G, Olsson K, Enfors S-O. Comparison of the baker's yeast process performance in laboratory and production scale. Bioprocess Biosyst Eng 1998;18: 135–42.
Gernaey KV, van Loosdrecht MCM, Henze M, Lind M, Jørgensen SB. Activated sludge wastewater treatment plant modelling and simulation: state of the art. Environ Modell Soft 2004;19:763–83.
Gernaey KV, Lantz AE, Tufvesson P, Woodley JM, Sin G. Application of mechanistic models to fermentation and biocatalysis for next-generation processes. Trends Biotechnol 2010;28:346–54.
Godin FB, Cooper DG, Rey AD. Development and solution of a cell mass population balance model applied to the SCF process. Chem Eng Sci 1999;54:565–78.
Gombert AK, Nielsen J. Mathematical modelling of metabolism. Curr Opin Biotechnol 2000;11:180–6.
Gunasekera TS, Sørensen A, Attfield PV, Sørensen SJ, Veal DA. Inducible Gene Expression by Nonculturable Bacteria in Milk after Pasteurization. Appl Environ Microbiol 2002;68:1988–93.
Gy P. Sampling for analytical purposes. John Wiley and Sons; 1998.
Hammes F, Egli T. Cytometric methods for measuring bacteria in water: advantages, pitfalls and applications. Anal Bioanal Chem 2010;397:1083–95.
Hanson GT, McAnaney TB, Park ES, Rendell MEP, Yarbrough DK, Chu S, et al. Green fluorescent protein variants as ratiometric dual emission pH sensors. 1. Structural characterization and preliminary application. Biochemistry 2002;41:15477–88.

Hao X, Wang Q, Zhang X, Cao Y, van Loosdrecht MCM. Experimental evaluation of decrease in bacterial activity due to cell death and activity decay in activated sludge. Water Res 2009;43:3604–12.
Hatzis C, Porro D. Morphologically-structured models of growing budding yeast populations. J Biotechnol 2006;124:420–38.
Hatzis C, Srienc F, Fredrickson AG. Multistaged corpuscular models of microbial growth: Monte Carlo simulations. Biosystems 1995;36:19–35.
Heard J, Johnson BB, Wells JD, Angove MJ. Measuring 'hydrophobicity' of filamentous bacteria found in wastewater treatment plants. Colloids Surf B 2009;72:289–94.
Heim R, Tsien RY. Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. Curr Biol 1996;6:178–82.
Heim R, Prasher DC, Tsien RY. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. PNAS 1994;91:12501–4.
Henson MA. Dynamic modeling of microbial cell populations. Curr Opin Biotechnol 2003;14:460–7.
Henson MA. Modeling the synchronization of yeast respiratory oscillations. J Theor Biol 2004;231:443–58.
Henson MA. Cell ensemble modeling of metabolic oscillations in continuous yeast cultures. Comput Chem Eng 2005;29:645–61.
Henson MA, Müller D, Reuss M. Cell population modelling of yeast glycolytic oscillations. Biochem J 2002;368:433–46.
Henze M, Gujer W, Mino T, van Loosedrecht M. Activated sludge models asm1, asm2, asm2d and asm3. IWA Publishing; 2000.
Hermelink A, Brauer A, Lasch P, Naumann D. Phenotypic heterogeneity within microbial populations at the single-cell level investigated by confocal Raman microspectroscopy. Analyst 2009;134:1149–53.
Hewitt CJ, Nebe-von-Caron G, Axelsson B, McFarlane CM, Nienow AW. Studies related to the scale-up of high-cell-density *E. coli* fed-batch fermentation using multiparamenter flow cytometry: effect of a changing microenvironment with respect to glucose and dissolved oxygen concentration. Biotechnol Bioeng 2000;70:381–90.
Hirschey MD, Han Y, Stucky GD, Butler A. Imaging *Escherichia coli* using functionalized core/shell cdse/cds quantum dots. J Biol Inorg Chem 2006;11:663–9.
Hjortsø M. Solution and properties of age population balance models which assume discrete division ages. J Biotechnol 1995;42:271–80.
Hjortsø MA, Bailey JE. Transient responses of budding yeast populations. Math Biosci 1983;63:121–48.
Hjortsø MA, Nielsen J. Population balance models of autonomous microbial oscillations. J Biotechnol 1995;42:255–69.
Hoefel D, Grooby WL, Monis PT, Andrews S, Saint CP. Enumeration of water-borne bacteria using viability assays and flow cytometry: a comparison to culture-based techniques. J Microbiol Methods 2003;55:585–97.
Holm-Nielsen JB, Dahl CK, Esbensen KH. Representative sampling for process analytical characterization of heterogeneous bioslurry systems — a reference study of sampling issues in pat. Chemometr Intell Lab 2006;83:114–26.
Hu C, Chinenov Y, Kerppola TK. Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. Mol Cell 2002;9:789–98.
Huang H, Ridgway D, Gu T, Moo-Young M. A segregated model for heterologous amylase production by bacillus subtilis. Enzyme Microb Tech 2003;32:407–13.
Ishii S, Tago K, Senoo K. Single-cell analysis and isolation for microbiology and biotechnology: methods and applications. Appl Microbiol Biotechnol 2010;86:1281–92.
Islam R, Tisi D, Levy M, Lye G. Scale-up of *Escherichia coli* growth and recombinant protein expression conditions from microwell to laboratory and pilot scale based on matched kLa. Biotechnol Bioeng 2008;99:1128–39.
Jares-Erijman EA, Jovin TM. Fret imaging. Nat Biotechnol 2003;21:1387–95.
Jares-Erijman EA, Jovin TM. Imaging molecular interactions in living cells by fret microscopy. Curr Opin Chem Biol 2006;10:409–16.
Jen C, Chou C, Hsu P, Yu S, Chen W, Lay J, et al. Flow-fish analysis and isolation of clostridial strains in an anaerobic semi-solid bio-hydrogen producing system by hydrogenase gene target. Appl Microbiol Biotechnol 2007;74:1126–34.
Jenne M, Reuss M. A critical assessment on the use of k-ε turbulence models for simulation of the turbulent liquid flow induced by a rushton-turbine in baffled stirred-tank reactors. Chem Eng Sci 1999;54:3921–41.
Jones KD, Kompala DS. Cybernetic model of the growth dynamics of *Saccharomyces cerevisiae* in batch and continuous cultures. J Biotechnol 1999;71:105–31.
Joung JK, Ramm EI, Pabo CO. A bacterial two-hybrid selection system for studying protein–DNA and protein–protein interactions. PNAS 2000;97:7382–7.
Joux F, Lebaron P. Use of fluorescent probes to assess physiological functions of bacteriaat single-cell level. Microbes Infect 2000;2:1523–35.
Kacmar J, Gilbert A, Cockrell J, Srienc F. The cytostat: a new way to study cell physiology in a precisely defined environment. J Biotechnol 2006;126:163–72.
Kahana JA, Schnapp BJ, Silver PA. Kinetics of spindle pole body separation in budding yeast. PNAS 1995;92:9707–11.
Kalyuzhnaya MG, Zabinsky R, Bowerman S, Baker DR, Lidstrom ME, Chistoserdova L. Fluorescence in situ hybridization-flow cytometry-cell sorting-based method for separation and enrichment of type I and type II methanotroph populations. Appl Environ Microbiol 2006;72:4293–301.
Kalyuzhnaya MG, Lidstrom ME, Chistoserdova L. Real-time detection of actively metabolizing microbes by redox sensing as applied to methylotroph populations in lake Washington. ISME J 2008;2:696–706.
Kavousanakis M, Mantzaris N, Boudouvis A. A novel free boundary algorithm for the solution of cell population balance models. Chem Eng Sci 2009;64:4247–61.
Kelly WJ. Using computational fluid dynamics to characterize and improve bioreactor performance. Biotechnol Appl Biochem 2008;49:225–38.
Kitano H. Cancer as a robust system: implications for anticancer therapy. Nat Rev Cancer 2004;4:227–35.
Kloeke FVO, Geesey G. Localization and identification of populations of phosphatase-active bacterial cells associated with activated sludge flocs. Microb Ecol 1999;38:201–14.
Kloepfer JA, Mielke RE, Wong MS, Nealson KH, Stucky G, Nadeau JL. Quantum dots as strain-and metabolism-specific microbiological labels. Appl Environ Microbiol 2003;69:4205–13.
Kloepfer JA, Mielke RE, Nadeau JL. Uptake of cdse and cdse/zns quantum dots into bacteria via purine-dependent mechanisms. Appl Environ Microb 2005;71:2548.
Kneen M, Farinas J, Li Y, Verkman A. Green fluorescent protein as a noninvasive intracellular ph indicator. Biophys J 1998;74:1591–9.
Krafft C, Dietzek B, Popp J. Raman and cars microspectroscopy of cells and tissues. Analyst 2009;134:1046–57.
Kragelund C, Nielsen JL, Thomsen TR, Nielsen PH. Ecophysiology of the filamentous alphaproteobacterium meganema perideroedes in activated sludge. FEMS Microbiol Ecol 2005;54:111–22.
Kurtz MJ, Zhu G, Zamamiri A, Henson MA, Hjortso MA. Control of oscillating microbial cultures described by population balance models. Ind Eng Chem Res 1998;37:4059–70.
Lacroix C, Yildirim S. Fermentation technologies for the production of biotics with high viability and functionality. Curr Opin Biotechnol 2007;18:176–83.
Lapin A, Muller D, Reuss M. Dynamic behavior of microbial populations in stirred bioreactors simulated with Euler–Lagrange methods: traveling along the lifelines of single cells. Ind Eng Chem Res 2004;43:4647–56.
Lapin A, Schmid J, Reuss M. Modeling the dynamics of e. coli populations in the three-dimensional turbulent field of a stirred-tank bioreactor — a structured-segregated approach. Chem Eng Sci 2006;61:4783–97.
Lara AR, Galindo E, Ramírez OT, Palomares LA. Living with heterogeneities in bioreactors. Mol Biotechnol 2006a;34:355–81.
Lara AR, Vazquez-Limón C, Gosset G, Bolívar F, López-Munguía A, Ramírez O. Engineering *Escherichia coli* to improve culture performance and reduce formation of by-product during recombinant protein product under transient intermittent anaerobic conditions. Biotechnol Bioeng 2006b;94:1164–75.
Lara AR, Taymaz-Nikerel H, Mashego MR, van Gulik WM, Heijnen JJ, Ramírez OT, et al. Fast dynamic response of the fermentative metabolism of *Escherichia coli* to aerobic and anaerobic glucose pulses. Biotechnol Bioeng 2009;104:1153–61.
Larsen P, Nielsen JL, Dueholm MS, Wetzel R, Otzen D, Nielsen PH. Amyloid adhesins are abundant in natural biofilms. Environ Microbiol 2007;9:3077–90.
Larsen P, Olesen BH, Nielsen PH, Nielsen JL. Quantification of lipids and protein in thin biofilms by fluorescence staining. Biofouling 2008;24:241–50.
Larsson G, Törnkvist M, Wernersson ES, Trägårdh C, Noorman H, Enfors S-O. Substrate gradients in bioreactors: origin and consequences. Bioprocess Biosyst Eng 1996;14:281–9.
Lee N, Nielsen PH, Andreasen KH, Juretschko S, Nielsen JL, Schleifer K, et al. Combination of fluorescent in situ hybridization and microautoradiography — a new tool for structure-function analyses in microbial ecology. Appl Environ Microbiol 1999;65:1289–97.
Lee HLT, Boccazzi P, Gorret N, Ram RJ, Sinskey AJ. In situ bioprocess monitoring of *Escherichia coli* bioreactions using Raman spectroscopy. Vib Spectrosc 2004;35:131–7.
Lee JA, Spidlen J, Boyce K, Cai J, Corsbie N, Dalphin M, et al. MIFlowCyt: the minimum information about a flow cytometry experiment. Cytometry A 2008;73A:926–30.
Lengeler JW, Drews G, Schlegel H, editors. Biology of prokaryotes. New York: Thieme; 1999.
Lenski RE, Travisano M. Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. PNAS 1994;91:6808–14.
Li JJ, Herskowitz I. Isolation of orc6, a component of the yeast origin recognition complex by a one-hybrid system. Science 1993;262:1870–4.
Link AJ, Jeong KJ, Georgiou G. Beyond toothpicks: new methods for isolating mutant bacteria. Nat Rev Microbiol 2007;5:680–8.
Liou J, Srienc F, Fredrickson AG. Solutions of population balance models based on a successive generations approach. Chem Eng Sci 1997;52:1529–40.
Lisle JT, Pyle BH, McFeters GA. The use of multiple indices of physiological activity to access viability in chlorine disinfected *Escherichia coli* o157:h7. Lett Appl Microbiol 1999;29:42–7.
Lo K, Brinkman RR, Gottardo R. Automated gating of flow cytometry data via robust model-based clustering. Cytometry A 2008;73A:321–32.
Luzyanina T, Roose D, Schenkel T, Sester M, Ehl S, Meyerhans A, et al. Numerical modeling of label-structured cell population growth using CFSE distribution data. Theor Biol Med Model 2007;4:26.
Lyons AB. Analysing cell division in vivo and in vitro using flow cytometric measurement of CFSE dye dilution. J Immunol Methods 2000;243:147–54.
Mandy FF, Bergeron M, Minkus T. Principles of flow cytometry. Transfus Sci 1995;16:303–14.
Mantzaris NV, Daoutidis P. Cell population balance modeling and control in continuous bioreactors. J Process Contr 2004;14:775–84.
Mantzaris NV, Liou J, Daoutidis P, Srienc F. Numerical solution of a mass structured cell population balance model in an environment of changing substrate concentration. J Biotechnol 1999;71:157–74.
Mantzaris NV, Daoutidis P, Srienc F. Numerical solution of multi-variable cell population balance models: I. Finite difference methods. Comput Chem Eng 2001a;25:1411–40.
Mantzaris NV, Daoutidis P, Srienc F. Numerical solution of multi-variable cell population balance models. III. Finite element methods. Comput Chem Eng 2001b;25:1463–81.
Mantzaris NV, Daoutidis P, Srienc F. Numerical solution of multi-variable cell population balance models. II. Spectral methods. Comput Chem Eng 2001c;25:1441–62.

Mantzaris NV, Srienc F, Daoutidis P. Nonlinear productivity control using a multi-staged cell population balance model. Chem Eng Sci 2002;57:1-14.
Mariani L, Martegani E, Alberghina L. Yeast population models for monitoring and control of biotechnical processes. IEE Proc 1986;133:210–6.
Marles-Wright J, Lewis RJ. Stress responses of bacteria. Curr Opin Struct Biol 2007;17: 755–60.
Mateus C, Avery SA. Destabilized green fluorescent protein for monitoring dynamic changes in yeast gene expression with flow cytometry. Yeast 2000;16:1313–23.
Mattiasson B, Håkanson H. Sampling and sample handling — crucial steps in process monitoring and control. Trends Biotechnol 1993;11:136–42.
Maukonen J, Mättö J, Wirtanen G, Raaska L, Mattila-Sandholm T, Saarela M. Methodologies for the characterization of microbes in industrial environments: a review. J Ind Microbiol Biotechnol 2003;30:327–56.
Mazumder S, Dey R, Mitra MK, Mukherjee S, Das GC. Review: biofunctionalized quantum dots in biology and medicine. J Nanomater 2009:1-17.
Mazumder S, Sarkar J, Dey R, Mitra MK, Mukherjee S, Das GC. Biofunctionalised quantum dots for sensing and identification of waterborne bacterial pathogens. J Exp Nanosci 2010;5:438–46.
Mhaskar P, Hjortsø MA, Henson MA. Cell population modeling and parameter estimation for continuous cultures of *Saccharomyces cerevisiae*. Biotechnol Prog 2002;18:1010–26.
Micale G, Montante G, Grisafi F, Brucato A, Godfrey J. CFD simulation of particle distribution in stirred vessels. Chem Eng Res Des 2000;78:435–44.
Michalet X, Pinaud FF, Bentolila LA, Tsay JM, Doose S, Li JJ, et al. Quantum dots for live cells, in vivo imaging, and diagnostics. Science 2005;307:538–44.
Micheletti M, Lye GJ. Microscale bioprocess optimisation. Curr Opin Biotech 2006;17: 611–8.
Miesenböck G, De Angelis DA, Rothman JE. Visualizing secretion and synaptic transmission with ph-sensitive green fluorescent proteins. Nature 1998;394:192–5.
Millard P, Roth B. Fluorescence-based methods for microbial characterization and viability assessment. Biotechnol Intl 1997;1:291–6.
Miller JH, Reznikoff WS, Silverstone AE, Ippen K, Signer ER, Beckwith JR. Fusions of the lac and trp regions of the *Escherichia coli* chromosome. J Bacteriol 1970;104: 1273–9.
Miyanaga K, Takano S, Morono Y, Hori K, Unno H, Tanji Y. Optimization of distinction between viable and dead cells by fluorescent staining method and its application to bacterial consortia. Biochem Eng J 2007;37:56–61.
Moat A, Foster JW, Spector MP. Microbial physiology. 4th ed. New York: John Wiley & Sons; 2002.
Mortensen PP, Bro R. Real-time monitoring and chemical profiling of a cultivation process. Chemometr Intell Lab 2006;84:106–13.
Müller S, Ullrich S, Lösche A, Loffhagen N, Babel W. Flow cytometric techniques to characterise physiological states of acinetobacter calcoaceticus. J Microbiol Methods 2000;40:67–77.
Müller S, Harms H, Bley T. Origin and analysis of microbial population heterogeneity in bioprocesses. Curr Opin Biotechnol 2010;21:100–13.
Nasipuri A, Subramanian KR, Ogunro V, Daniels JL, Hilger HA. Development of a wireless sensor network for monitoring a bioreactor landfill. ASCE Conf Proc 2006;187:10.
Nebe-von-Caron G. Standardization in microbial cytometry. Cytometry A 2009;75A: 86–9.
Nebe-von-Caron G, Stephens PJ, Hewitt CJ, Powell JR, Badley RA. Analysis of bacterial function by multi-colour fluorescence flow cytometry and single cell sorting. J Microbiol Methods 2000;42:97-114.
Neu TR, Manz B, Volke F, Dynes JJ, Hitchcock AP, Lawrence JR. Advanced imaging techniques for assessment of structure, composition and function in biofilm systems. FEMS Microbiol Ecol 2010;72:1-21.
Niedenthal RK, Riles L, Johnston M, Hegemann JH. Green fluorescent protein as a marker for gene expression and subcellular localization in budding yeast. Yeast 1996;12:773–86.
Nielsen JL, Nielsen PH. Advances in microscopy: microautoradiography of single cells. Methods Enzymol 2005;397:237–56.
Nielsen TH, Sjøholm OR. Sørensen JS Multiple physiological states of a *Pseudomonas fluorescens* dr54 biocontrol inoculant monitored by a new flow cytometry protocol. FEMS Microbiol Ecol 2009;67:479–90.
Nielsen J, Villadsen J. Modelling of microbial kinetics. Chem Eng Sci 1992;47:4225–70.
Nielsen JL, Mikkelsen LH, Nielsen PH. In situ detection of cell surface hydrophobicity of probe-defined bacteria in activated sludge. Water Sci Technol 2001;43:97-103.
Nielsen JL, Aquino de Muro M, Nielsen PH. Evaluation of the redox dye 5-cyano-2, 3-tolyl-tetrazolium chloride for activity studies by simultaneous use of micro-autoradiography and fluorescence in situ hybridization. Appl Environ Microbiol 2003a;69:641–3.
Nielsen JL, Christensen D, Kloppenborg M, Nielsen PH. Quantification of cell-specific substrate uptake by probe-defined bacteria under in situ conditions by micro-autoradiography and fluorescence in situ hybridization. Environ Microbiol 2003b;5:202–11.
Nocker A, Cheung C, Camper AK. Comparison of propidium monoazide with ethidium monoazide for differentiation of live vs. dead bacteria by selective removal of DNA from dead cells. J Microbiol Methods 2006;67:310–20.
Norman A, Hansen LH, Sørensen SJ. Construction of an SOS-GFP whole-cell biosensor using the ColD *cda*-Promoter with higher sensitivity towards genotoxic compounds than constructs based on *recA*-, *umuDC*- or *sulA*-promoters. Appl Environ Microbiol 2005;71:2338–46.
Norman A, Hansen LH, Sørensen SJ. A flow cytometry-optimized assay using an sos-green fluorescent protein (sos-GFP) whole-cell biosensor for the detection of genotoxins in complex environments. Mutat Res-Envir Muta 2006;603:164–72.
Novo D, Perlmutter NG, Hunt RH, Shapiro HM. Accurate flow cytometric membrane potential measurement in bacteria using diethyloxacarbocyanine and a ratiometric technique. Cytometry 1999;35:55–63.
Oerther DB, Pernthaler J, Schramm A, Amann R, Raskin L. Monitoring precursor 16s rRNAs of *Acinetobacter spp.* in activated sludge wastewater treatment systems. Appl Environ Microbiol 2000;66:2154–65.
Onyeaka H, Nienow AW, Hewitt CJ. Further studies related to the scale-up of high cell density *Escherichia coli* fed-batch fermentations. Biotechnol Bioeng 2003;84: 474–84.
Ostle AG, Holt JG. Nile blue a as a fluorescent stain for poly-beta-hydroxybutyrate. Appl Environ Microbiol 1982;44:238–41.
Page WJ, Tenove CJ. Quantitation of poly-β-hydroxybutyrate by fluorescence of bacteria and granules stained with Nile blue A. Biotechnol Tech 1996;10:215–20.
Papadimitriou K, Pratsinis H, Nebe-von-Caron G, Kletsas D, Tsakalidou E. Acid tolerance of streptococcus macedonicus as assessed by flow cytometry and single-cell sorting. Appl Environ Microbiol 2007;73:465–76.
Papagianni M. Fungal morphology and metabolite production in submerged mycelial processes. Biotechnol Adv 2004;22:189–259.
Patterson GH, Knobel SM, Sharif WD, Kain SR, Piston DW. Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. Biophys J 1997;73:2782–90.
Pfleger KDG, Eidne KA. Illuminating insights into protein–protein interactions using bioluminescence resonance energy transfer (BRET). Nat Methods 2006;3:165–74.
Pollard DJ, Ison AP, Shamlou PA, Lilly MD. Reactor heterogeneity with *Saccharopolyspora erythraea* airlift fermentations. Biotechnol Bioeng 1998;58:453–63.
Ponciano JM, La H, Joyce P, Forney LJ. Evolution of diversity in spatially structured *Escherichia coli* populations. Appl Environ Microbiol 2009;75:6047–54.
Porro D, Vai M, Vanoni M, Alberghina L, Hatzis C. Analysis and modeling of growing budding yeast populations at the single cell level. Cytometry A 2009;75:114–20.
Poulsen LK, Ballard G, Stahl DA. Use of rRNA fluorescence in situ hybridization for measuring the activity of single cells in young and established biofilms. Appl Environ Microbiol 1993;59:1354–60.
Prinz A, Reither G, Diskar M, Schultz C. Fluorescence and bioluminescence procedures for functional proteomics. Proteomics 2008;8:1179–96.
Pyle B, Broadaway S, McFeters G. Factors affecting the determination of respiratory activity on the basis of cyanoditolyl tetrazolium chloride reduction with membrane filtration. Appl Environ Microbiol 1995;61:4304–9.
Pyne S, Hu X, Wang K, Rossin E, Lin T-I, Maier LM, et al. Automated high-dimensional flow cytometric data analysis. Proc Natl Acad Sci U S A 2009;106:8519–24.
Rademaker J, de Bruijn FJ. Characterization and classification of microbes by rep-pcr genomic fingerprinting and computer-assisted pattern analysis. In: Caetano-Anolle G, Gresshoff PM, editors. DNA Markers: Protocols, Applications and Overviews. New York: John Wiley; 1997. p. 151–71.
Ramkrishna D. Toward a self-similar theory of microbial populations. Biotechnol Bioeng 1994;43:138–48.
Ramkrishna D. Population balances: theory and applications to particulate systems in engineering. 1st ed. Academic Press; 2000.
Ramkrishna D. On modeling of bioreactors for control. J Process Contr 2003;13: 581–9.
Rault A, Bouix M, Beal C. Fermentation ph influences the physiological-state dynamics of lactobacillus bulgaricus cfl1 during ph-controlled culture. Appl Environ Microbiol 2009;75:4374–81.
Regan JM, Oldenburg PS, Park HD, Harrington GW, Noguera DR. Simultaneous determination of bacterial viability and identity in biofilms using ethidium monoazide and fluorescent in situ hybridization. Water Sci Technol 2003;47: 123–8.
Revstedt J, Fuchs L, Trägardh C. Large eddy simulations of the turbulent flow in a stirred reactor. Chem Eng Sci 1998;53:4041–53.
Riedy MC, Muirhead KA, Jensen CP, Stewart CC. Use of a photolabeling technique to identify nonviable cells in fixed homologous or heterologous cell populations. Cytometry 1991;12:133–9.
Rieseberg M, Kasper C, Reardon KF, Scheper T. Flow cytometry in biotechnology. Appl Microbiol Biotechnol 2001;56:350–60.
Ritzka A, Sosnitza P, Ulber R, Scheper T. Fermentation monitoring and process control. Curr Opin Biotechnol 1997;8:160–4.
Rodriguez GG, Phipps D, Ishiguro K, Ridgway HF. Use of a fluorescent redox probe for direct visualization of actively respiring bacteria. Appl Environ Microbiol 1992;58: 1801–8.
Ron E. Bacterial stress response. The Prokaryotes. 3rd edition. New York: Springer; 2006. p. 1012–27.
Rösch P, Harz M, Schmitt M, Peschke K, Ronneberger O, Burkhardt H, et al. Chemotaxonomic identification of single bacteria by micro-Raman spectroscopy: application to clean-room-relevant biological contaminations. Appl Environ Microbiol 2005a;71:1626–37.
Rösch P, Harz M, Schmitt M, Popp J. Raman spectroscopic identification of single yeast cells. J Raman Spectrosc 2005b;36:377–9.
Saint-Ruf C, Cordier C, Megret J, Matic I. Reliable detection of dead microbial cells by using fluorescent hydrazides. Appl Environ Microbiol 2010;76:1674–8.
Schäpper D, Alam M, Szita N, Eliasson Lantz A, Gernaey KV. Application of microbioreactors in fermentation process development: a review. Anal Bioanal Chem 2009;395:679–95.
Schmid M, Schmitz-Esser S, Jetten M, Wagner M. 16s–23s rDNA intergenic spacer and 23s rDNA of anaerobic ammonium-oxidizing bacteria: implications for phylogeny and in situ detection. Environ Microbiol 2001;3:450–9.
Schmid A, Kortmann H, Dittrich PS, Blank LM. Chemical and biological single cell analysis. Curr Opin Biotechnol 2010;21:12–20.

Schneider R, Wolpert C, Guilloteau H, Balan L, Lambert J, Merlin C. The exposure of bacteria to cdte-core quantum dots: the importance of surface chemistry on cytotoxicity. Nanotechnology 2009;20:225101.
Schuster KC, Urlaub E, Gapes JR. Single-cell analysis of bacteria by Raman microscopy: spectral information on the chemical composition of cells and on the heterogeneity in a culture. J Microbiol Methods 2000a;42:29–38.
Schuster KC, Reese I, Urlaub E, Gapes JR, Lendl B. Multidimensional information on the chemical composition of single bacterial cells by confocal Raman microspectroscopy. Anal Chem 2000b;72:5529–34.
Serrazanetti DI, Guerzoni ME, Corsetti A, Vogel R. Metabolic impact and potential exploitation of the stress reactions in lactobacilli. Food Microbiol 2009;26:700–11.
Shapiro HM. Microbial analysis at the single-cell level: tasks and techniques. J Microbiol Methods 2000;42:3-16.
Shapiro HM. The evolution of cytometers. Cytometry A 2004;58A:13–20.
Sherer E, Tocce E, Hannemann R, Rundell A, Ramkrishna D. Identification of age-structured models: cell cycle phase transitions. Biotechnol Bioeng 2008;99:960–74.
Sherr BF, del Giorgio P, Sherr EB. Estimating abundance and single-cell characteristics of respiring bacteria via the redox dye ctc. Aquat Microb Ecol 1999;18:117–31.
Shi L, Günther S, Hübschmann T, Wick LY, Harms H, Müller S. Limits of propidium iodide as a cell viability indicator for environmental bacteria. Cytometry A 2007;71:592–8.
Silhavy TJ, Beckwith JR. Uses of lac fusions for the study of biological problems. Microbiol Rev 1985;49:398–418.
Smith JJ, McFeters GA. Mechanisms of int (2-(4-iodophenyl)-3-(4-nitrophenyl)-5-phenyl tetrazolium chloride), and ctc (5-cyano-2,3-ditolyl tetrazolium chloride) reduction in *Escherichia coli* k-12. J Microbiol Methods 1997;29:161–75.
Sørensen SJ, Sørensen AH, Hansen LH, Oregaard G, Veal D. Direct detection and quantification of horizontal gene transfer by using flow cytometry and GFP as a reporter gene. Curr Microbiol 2003;47:0129–33.
Sørensen SJ, Burmølle M, Hansen LH. Making bio-sense of toxicity: new developments in whole-cell biosensors. Curr Opin Biotechnol 2006;17:11–6.
Stagljar I, Korostensky C, Johnsson N, te Heesen S. A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins in vivo. PNAS 1998;95:5187–92.
Stamatakis M. Cell population balance, ensemble and continuum modeling frameworks: conditional equivalence and hybrid approaches. Chem Eng Sci 2010;65:1008–15.
Stelling J. Mathematical models in microbial systems biology. Curr Opin Microbiol 2004;7:513–8.
Stewart PS, Franklin MJ. Physiological heterogeneity in biofilms. Nat Rev Microbiol 2008;6:199–210.
Subramanian G, Ramkrishna D, Fredrickson A, Tsuchiya H. On the mass distribution model for microbial cell populations. B Math Biol 1970;32:521–37.
Sumner ER, Avery SV. Phenotypic heterogeneity: differential stress resistance among individual cells of the yeast *Saccharomyces cerevisiae*. Microbiology 2002;148:345–51 M.
Sweere APJ, Giesselbach J, Barendse R, Krieger R, Honderd G, Luyben KCAM. Modelling the dynamic behaviour of *Saccharomyces cerevisiae* and its application in control experiments. Appl Microbiol Biotechnol 1988a;28:116–27.
Sweere APJ, Mesters JR, Janse L, Luyben KCAM, Kossen NWF. Experimental simulation of oxygen profiles and their influence on baker's yeast production: I. One-fermentor system. Biotechnol Bioeng 1988b;31:567–78.
Szita N, Polizzi K, Jaccard N, Baganz F. Microfluidic approaches for systems and synthetic biology. Curr Opin Biotechnol 2010;21:517–23.
Takamatsu T, Shioya S, Chikatani H, Dairaku K. Comparison of simple population models in a baker's yeast fed-batch culture. Chem Eng Sci 1985;40:499–507.
Teusink B, Smid EJ. Modelling strategies for the industrial exploitation of lactic acid bacteria. Nat Rev Microbiol 2006;4:46–56.
Theobald U, Mailinger W, Baltes M, Rizzi M, Reuss M. In vivo analysis of metabolic dynamics in *Saccharomyces cerevisiae*: I. Experimental observations. Biotechnol Bioeng 1997;55:305–16.
Tracy BP, Gaida SM, Papoutsakis ET. Development and application of flow-cytometric techniques for analyzing and sorting endospore-forming clostridia. Appl Environ Microbiol 2008;74:7497–506.
Tracy BP, Gaida SM, Papoutsakis ET. Flow cytometry for bacteria: enabling metabolic engineering, synthetic biology and the elucidation of complex phenotypes. Curr Opin Biotechnol 2010;21:85–99.
Tsien RY. Fluorescent probes of cell signaling. Annu Rev Neurosci 1989;12:227–53.
Tsien RY. The green fluorescent protein. Annu Rev Biochem 1998;67:509–44.
Ueckert JE, Nebe-von-Caron G, Bos AP, ter Steeg PF. Flow cytometric analysis of *Lactobacillum platrum* to monitor lag times, cell division and injury. Lett Appl Microbiol 1997;25:295–9.
Ulber R, Frerichs J, Beutel S. Optical sensor systems for bioprocess monitoring. Anal Bioanal Chem 2003;376:342–8.
Veal DA, Deere D, Ferrari B, Piper J, Attfield PV. Fluorescence staining and flow cytometry for monitoring microbial cells. J Immunol Methods 2000;243:191–210.
Veening J, Smits WK, Hamoen LW, Jongbloed JDH, Kuipers OP. Visualization of differential gene expression by improved cyan fluorescent protein and yellow fluorescent protein production in bacillus subtilis. Appl Environ Microbiol 2004;70:6809–15.
Wagner M, Horn M, Daims H. Fluorescence in situ hybridisation for the identification and characterisation of prokaryotes. Curr Opin Microbiol 2003;6:302–9.
Waldo GS, Standish BM, Berendzen J, Terwilliger TC. Rapid protein-folding assay using green fluorescent protein. Nat Biotechnol 1999;17:691–5.
Walling MA, Novak JA, Shepard JRE. Quantum dots for live cell and in vivo imaging. Int J Mol Sci 2009;10:441–91.
Waltermann M, Steinbuchel A. Neutral lipid bodies in prokaryotes: recent insights into structure, formation, and relationship to eukaryotic lipid depots. J Bacteriol 2005;187:3607–19.
Ward WW. Properties of the coelenterate green-fluorescent proteins. Bioluminescence and Chemiluminescence. New York: Academic Press; 1981. p. 235–42.
Ward WW, Bokman SH. Reversible denaturation of aequorea green-fluorescent protein: physical separation and characterization of the renatured protein. Biochemistry 1982;21:4535–40.
Wheals AE. Size control models of *Saccharomyces cerevisiae* cell proliferation. Mol Cell Biol 1982;2:361–8.
Whoriskey SK, Nghiem VH, Leong PM, Masson JM, Miller JH. Genetic rearrangements and gene amplification in *Escherichia coli*: DNA sequences at the junctures of amplified gene fusions. Genes Dev 1987;1:227–37.
Wick L, Egli T. Molecular components of physiological stress responses in *Escherichia coli*. Physiological Stress Responses in Bioprocesses. New York: Springer; 2004. p. 1-45.
Wigley WC, Stidham RD, Smith NM, Hunt JF, Thomas PJ. Protein solubility and folding monitored in vivo by structural complementation of a genetic marker protein. Nat Biotechnol 2001;19:131–6.
Wimpenny J, Manz W, Szewzyk U. Heterogeneity in biofilms. FEMS Microbiol Rev 2000;24:661–71.
Wu L, Kellogg L, Devol AH, Tiedje JM, Zhou J. Microarray-based characterization of microbial community functional structure and heterogeneity in marine sediments from the gulf of Mexico. Appl Environ Microbiol 2008;74:4516–29.
Wu SM, Tian ZQ, Zhang ZL, Huang BH, Jiang P, Xie ZX, et al. Direct fluorescence *in situ* hybridization (FISH) in *Escherichia coli* with a target-specific quantum dot-based molecular beacon. Biosens Bioelectron 2010;26:491–6.
Wuertz S, Bishop PL, Wilderer PA. Biofilms in wastewater treatment: an interdisciplinary approach. IWA Publishing; 2003.
Yang T, Cheng L, Kain S. Optimized codon usage and chromophore mutations provide enhanced sensitivity with the green fluorescent protein. Nucleic Acids Res 1996;24:4592–3.
Ye RW, Wang T, Bedzyk L, Croker KM. Applications of DNA microarrays in microbial systems. J Microbiol Methods 2001;47:257–72.
Yen HS, Xu Q, Chou DM, Zhao Z, Elledge SJ. Global protein stability profiling in mammalian cells. Science 2008;322:918–23.
Young J, Henne K, Morgan J, Konopka A, Ramkrishna D. Cybernetic modeling of metabolism: towards a framework for rational design of recombinant organisms. Chem Eng Sci 2004;59:5041–9.
Zamamiri AM, Zhang Y, Henson MA, Hjortsø MA. Dynamics analysis of an age distribution model of oscillating yeast cultures. Chem Eng Sci 2002;57:2169–81.
Zhang B, Kraemer B, SenGupta D, Fields S, Wickens M. Yeast three-hybrid system to detect and analyze interactions between RNA and protein. Methods Enzymol 1999;306:93-113.
Zhang H, Zhang K, Fan S. CFD simulation coupled with population balance equations for aerated stirred bioreactors. Eng Life Sci 2009;9:421–30.
Zhao X, Hilliard LR, Mechery SJ, Wangt Y, Bagwe RP, Jint S, et al. Rapid bioassay for single bacterial cell quantitation using bioconjugated nanoparticles. Proc Natl Acad Sci U S A 2004;101:15027–32.
Zhu G, Zamamiri A, Henson MA, Hjortsø MA. Model predictive control of continuous yeast bioreactors using cell population balance models. Chem Eng Sci 2000;2000(55):6155–67.
Zita A, Hermansson M. Determination of bacterial cell surface hydrophobicity of single cells in cultures and in wastewater in situ. FEMS Microbiol Lett 1997;152:299–306.
Zotta T, Parente E, Ricciardi A. Viability staining and detection of metabolic activity of sourdough lactic acid bacteria under stress conditions. World J Microb Biot 2009;25:1119–24.

ARTICLE

BIOTECHNOLOGY and BIOENGINEERING

Cell Mass and Cell Cycle Dynamics of an Asynchronous Budding Yeast Population: Experimental Observations, Flow Cytometry Data Analysis, and Multi-Scale Modeling

Rita Lencastre Fernandes,[1] Magnus Carlquist,[2,3] Luisa Lundin,[4] Anna-Lena Heins,[2] Abhishek Dutta,[5] Søren J. Sørensen,[4] Anker D. Jensen,[6] Ingmar Nopens,[5] Anna Eliasson Lantz,[2] Krist V. Gernaey[1]

[1]Center for Process Engineering and Technology, Department of Chemical and Biochemical Engineering, Technical University of Denmark, DK-2800, Kongens Lyngby, Denmark; telephone: +45 45 25 29 70; fax: +45 45 93 29 06; e-mail: kvg@kt.dtu.dk

[2]Center for Microbial Biotechnology, Department of Systems Biology, Technical University of Denmark, Kongens Lyngby, Denmark

[3]Division of Applied Microbiology, Department of Chemistry, Lund University, Lund, Sweden

[4]Molecular Microbial Ecology Group, Department of Biology, University of Copenhagen, Copenhagen, Denmark

[5]BIOMATH, Department of Mathematical Modelling, Statistics and Bioinformatics, Ghent University, Ghent, Belgium

[6]Center for Combustion and Harmful Emission Control, Department of Chemical and Biochemical Engineering, Technical University of Denmark, Kongens Lyngby, Denmark

ABSTRACT: Despite traditionally regarded as identical, cells in a microbial cultivation present a distribution of phenotypic traits, forming a heterogeneous cell population. Moreover, the degree of heterogeneity is notably enhanced by changes in micro-environmental conditions. A major development in experimental single-cell studies has taken place in the last decades. It has however not been fully accompanied by similar contributions within data analysis and mathematical modeling. Indeed, literature reporting, for example, quantitative analyses of experimental single-cell observations and validation of model predictions for cell property distributions against experimental data is scarce. This study focuses on the experimental and mathematical description of the dynamics of cell size and cell cycle position distributions, of a population of *Saccharomyces cerevisiae*, in response to the substrate consumption observed during batch cultivation. The good agreement between the proposed multi-scale model (a population balance model [PBM] coupled to an unstructured model) and experimental data (both the overall physiology and cell size and cell cycle distributions) indicates that a mechanistic model is a suitable tool for describing the microbial population dynamics in a bioreactor. This study therefore contributes towards the understanding of the development of heterogeneous populations during microbial cultivations. More generally, it consists of a step towards a paradigm change in the study and description of cell cultivations, where average cell behaviors observed experimentally now are interpreted as a potential joint result of various co-existing single-cell behaviors, rather than a unique response common to all cells in the cultivation.

Biotechnol. Bioeng. 2012;xxx: xxx–xxx.

KEYWORDS: population balance model (PBM); multiscale modeling; flow cytometry; standardized data analysis; *Saccharomyces cerevisiae*; total protein content; cell cycle

Correspondence to: K.V. Gernaey
Contract grant sponsor: Danish Council for Strategic Research
Contract grant number: 09-065160
Contract grant sponsor: ERA-IB (ERA-NET Industrial Biotechnology)
Contract grant number: EIB.08.031
Additional supporting information may be found in the online version of this article.
Received 24 August 2012; Accepted 5 October 2012
Accepted manuscript online xx Month 2012;
Article first published online in Wiley Online Library (wileyonlinelibrary.wiley.com).
DOI 10.1002/bit.24749

Introduction

Cells in a cultivation are traditionally regarded as identical, though microbial populations are most often heterogeneous. As such, distributions of phenotypic traits such as cell

size, enzymatic activities, and growth rate (Müller et al., 2010), which are often essential for fitness and development of the cells (Avery, 2006), are observed. This type of heterogeneity may originate from stochastic gene transcription, translation and regulation, differences in progression through cell cycle phases, and age distributions due to unequal partitioning upon division (Davey and Kell, 1996; Müller et al., 2010).

Frequently, the contribution of different aspects to the development of a heterogeneous population is difficult to distinguish. For example, in the case of *Saccharomyces cerevisiae*, cell viability after environmental stress (e.g., heat, copper, freeze–thaw) has shown to be dependent on the cell cycle phase (Carlquist et al., 2012; Howlett and Avery, 1999; Sumner et al., 2003), but also to be related to cell age (Kale and Jazwinski, 1996). Furthermore, the existence of different microenvironments within a reactor (i.e., spatial heterogeneity) may imply differential responses of cells as they experience a changing extracellular environment in their various trajectories throughout the reactor. Indeed, this fact has been pointed out as the main underlying cause for differences between cultivations performed in well-mixed lab scale bioreactors and more poorly mixed large-scale reactors (Enfors et al., 2001).

It is, therefore, in the understanding and description of the interplay between single cell response and the changing environment that the key to build improved predictive models lies (Lencastre Fernandes et al., 2011). In order to describe the dynamics of a microbial population in a standardized and quantitative fashion, an adequate set of cell properties has to be selected, and experimental observations have to be conducted and analyzed in a systematic way.

With regard to cell properties, cell size has often been used to describe budding yeast populations under various growth conditions (Porro et al., 2009). The choice of cell size relies on its tight coupling to cell growth and division. Indeed, cell size is a key feature affecting cellular design, fitness and function (Jorgensen and Tyers, 2004), and this is a reflection of the cellular capability of adjusting its growth rate to nutritional availability (Jorgensen et al., 2002; Saldanha et al., 2004). The regulation of growth ensures that cells attain a critical size before initiating the division process (Porro et al., 2003; Rupeš, 2002). In the particular case of *S. cerevisiae*, two critical sizes corresponding to the regulation points START (committing to budding, or budding transition) and division have been identified (as reviewed by Rupeš, 2002). Hence, using cell size as population physiological state descriptor allows for describing the distribution of cellular states. Experimentally, the distribution of cell size of a population is easily measured by using flow cytometry. In particular, the total protein content has been used as a reliable measure of cell size (Alberghina et al., 1998).

Cell size distributions of *S. cerevisiae* populations during balanced growth on various limiting substrates, as well as for various dilution rates (i.e., growth rates) have been reported and compared in different studies (e.g., Alberghina et al., 1998; Porro and Srienc, 1995; Porro et al., 2009; Vanoni et al., 1983). Larger critical cell sizes (at budding and, consequently, division) have previously been reported for higher growth rates, for example, during exponential growth on glucose relatively to ethanol (Alberghina et al., 1998; Cipollina et al., 2007), or with increasing dilution rates in glucose-limited continuous cultivations (Porro et al., 2003). Complementary information on the distribution of cells in cell cycle phases can be collected by measuring DNA distributions, yielding a better description of the cellular state (Porro et al., 2003). Also age dependency of the critical size upon the budding transition has been evaluated, as time spent in the G1 phase decreases with the number of cell cycles a mother cell has undertaken (Vanoni et al., 1983).

Population balance models (PBM) allow for a mathematical description of distributed cell properties within microbial populations (Fredrickson, 2003; Fredrickson et al., 1970). In previously published literature on PBM for microbial populations (Fredrickson, 2003; Hatzis et al., 1995; Mantzaris et al., 1999), cell size was used as model variable.

Hatzis and Porro (2006) proposed a multi-stage PBM accounting for non-budding and budding stages and continuous distributions of cell mass. Additionally, the model distinguished different generations, acknowledging the fact that the critical division size of an individual cell will increase for every cell cycle the cell undergoes. Although the formulation of this model (Hatzis and Porro, 2006) offers the possibility of including the dependence of the critical budding and division sizes on the substrate, this dependence has not been explicitly described neither have simulations under varying substrate conditions been reported. In previous work by Mantzaris et al. (1999), different substrate dependent growth kernels were tested and compared for a cell mass structured PBM. Validation of the assumptions taken or confrontation of model predictions with experimental data was however not reported.

PBM have also been used for describing mammalian cell systems (Fadda et al., 2012a,b; Karra et al., 2010; Liu et al., 2007; Sidoli et al., 2006). In these examples, model were based on kernel functions, and in some cases parameter values, that were similar to the ones initially proposed in theoretical studies on microbial populations (Mantzaris et al., 1999, 2001, 2002). Parameter estimation from experimental data have been reported by Mancuso et al. (2009) where two parameters corresponding to the maximum rate of cell growth and a power law order for a geometrical factor were tuned in order to minimize the differences between the experimental and predicted cell size distributions. Fadda et al. (2012b) qualitatively compared model predictions to experimental observations previously reported in the literature. Common to all these cases is the use of parameters and/or assumptions previously reported on theoretical work on PBM for microbial populations, for example, the partition distribution shape parameter. A comparison between models for microbial and mammalian cell populations and a discussion of the common features

and eventual assumptions specific to each type of cells has to our knowledge never been undertaken. Although of significant interest, such comparison falls out of the scope of this article.

This work aims at understanding and describing the interplay between single cell response, population dynamics and the changing environment in a systematic way so experimental observations can be translated into a mathematical model. Indeed, for the first time, to our knowledge, experimental evidences are used for defining the PBM model kernels as functions of the extracellular environment.

A PBM describing the cell size distributions for the non-budding and budding populations, during the growth of a population of *S. cerevisiae* along a glucose-limited batch cultivation. Experimentally, the development of an asynchronous population of *S. cerevisiae* was monitored during the different growth phases, with particular focus on the diauxic shift transition. General trends for cell size (total protein content) and cell cycle phase (DNA) distributions along the cultivation are reported and discussed. Furthermore, a standardized procedure for treatment of the gathered flow cytometric data was established for isolating subpopulations with high content of cells initiating the budding process and preparing for division. The trends identified based on these procedures were used in the definition of the critical budding and division sizes as function of the glucose and ethanol uptake rates, which are essential to the development of a PBM. Additionally, a simple unstructured model was coupled to the PBM in order to evaluate predictions of key changes in the composition of the extracellular environment (i.e., cultivation medium): the consumption of glucose, production and consumption of ethanol, as well as supply and consumption of dissolved oxygen.

Materials and Methods

Strain, Preculture and Batch Cultivations

For this study, batch cultivations of a haploid *S. cerevisiae* strain (prototrophic CEN.PK 113-5D) were performed in Braun Biostat 2 l bioreactors (B. Braun Biotech International, GmbH, Melsungen, Germany) with a working volume of 1.5 l. A defined mineral medium supplemented with 5 g L^{-1} glucose was used (Verduyn et al., 1992). Cultivation conditions were set to the following; aeration 1 vvm; temperature 30°C; stirring 600 rpm and pH 5.0 (automatically controlled by addition of 3.0 M KOH). Samples for OD_{600}, high performance liquid chromatography (HPLC) and flow cytometry analysis were taken approximately every 1 h, or every 30 min during the diauxic shift and early growth on ethanol. Further details on the strain, inocula preparation and sample treatment are provided as Supplementary Information (S1).

Single-Cell Analysis

A BD FACSAria III (Becton–Dickinson, Franklin Lakes, NJ) flow cytometer was used for single-cell analysis. For the simultaneous determination of total protein and DNA content, cells were stained with fluorescein isothiocyanate (FITC) and propidium iodide (PI; Sigma–Aldrich, Brøndby, Denmark) as described previously (Porro et al., 2003). Further details on the flow cytometer settings are provided as Supplementary Information (S1).

Computational Data Treatment

In this work, a systematic approach was applied to the flow cytometry data analysis: standardized procedures were developed for estimation of the critical budding and division sizes based on the experimental total protein and DNA content distributions. Firstly, a procedure was implemented in order to isolate a subpopulation with a high fraction of cells transitioning from G1 to S-phase, enabling estimation of the critical budding size. It relies on isolating the subpopulation presenting an intermediate DNA content (i.e., between 1 copy (1C) and 2 copies (2C) of the chromosome) and thus contained in the interpeak region in the DNA histogram, applying—but then in a more standardized fashion—the same principle used by Porro et al. (2003) for a manual gating approach. A second standardized procedure was developed in order to estimate the critical division size, based on the standard deviation of the 2C peak in the DNA histogram. The critical budding and division DNA bands were defined around the channel number at one standard deviation distance from the peak mode (Fig. 1). A band width of 10 channel numbers was defined in order to ensure a number of cells in the subpopulation of approximately 500. The critical budding and division cell sizes were defined as the mean total protein content of the cells belonging to the corresponding DNA critical bands. The budding index (BI), that is, the fraction of cells that have a bud, was estimated using the critical budding DNA band as threshold. Schematic representations of these procedures are provided as Supplementary Information (S2). Moreover, a comparison between the proposed BI estimation procedure based on flow cytometric data and the traditional microscopic counting is additionally provided as Supplementary Information (S2).

The robustness of the procedures was assessed by varying the band widths (data not shown). Such variation did not yield a significant effect on the results reported in this work, and the same correlations between critical sizes and substrate availability were observed. Although, simple threshold based methods as the ones proposed imply a certain degree of subjectivity, the aim was to develop simple decision algorithms that allow for treating all samples in the same fashion rather than relying on visual inspection and manual gating (often used in flow cytometric studies).

Processing and analysis of flow cytometry raw data was performed by using MatLab® R2009b (The MathWorks,

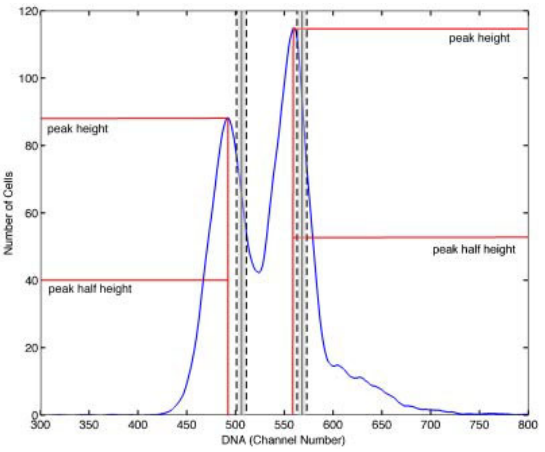

Figure 1. Estimation of the critical budding and division sizes based on the DNA histograms (linear scale) and the critical budding and division bands. Two peaks can be observed in the DNA histogram corresponding to the population with 1 copy of the chromosome (1C)—non-budding—to the left and the population with 2 copies (2C) to the right. Assuming the two peaks are normally distributed, the standard deviation σ_1 is estimated based on the 1C peak height ($\gamma(\mu_1) = h_{\max,1C}$), corresponding to the mean of the 1C peak distribution (μ_1). The critical budding band (gray vertical bar to the left) corresponds to the ± 5 ch. no. interval around $\nu_B = \mu_1 + \sigma_1$. Similarly, the critical division band (gray vertical bar to the right) is defined by an interval of ± 5 ch. no. around $\nu_D = \mu_2 + \sigma_2$, where μ_2 and σ_2 are the mean and standard deviation for the 2C peak. The critical budding and division sizes are defined as the mean total protein content of the cells pertaining to the respective critical bands.

Inc., Natick, MA). The measurement files, exported by the flow cytomer FACSAria II, were imported into MatLab®, using a "fcs data reader" routine (by L. Balkay, University of Debrecen, Hungary), freely available on the MatLab® File Exchange website.

Modeling Aspects

Population Balance Model

A two-stage PBM using cell total protein content (a measure of cell size) as model variable was developed. Channel number (ch. no.) was used as arbitrary unit for the cell total protein content. Further details on the correlation of channel number to cell size are provided as Supplementary Information (S3).

The PBM consists of two integro-partial differential equations—the population balance equations (PBE) for the non-budding and budding stages (Equations 1 and 2)—and the corresponding boundary and initial conditions (Equations 3–5). $N^{NB}(m, t)\mathrm{d}m$ and $N^{B}(m, t)\mathrm{d}m$ define the number of cells at time t with mass within the interval $[m, m + \mathrm{d}m]$, for the non-budding and budding stages respectively. Z describes the extracellular environment (i.e., the concentration of glucose, ethanol, and oxygen) at a given time point. A nomenclature list with a description and units for the model variables as well as a parameter list are provided as Supplementary Information (S4).

The left hand side of the PBEs (Equations 1 and 2) describes the accumulation of the number of cells in each stage and the growth of cells (i.e., continuous increase of the total cell protein content). In the case of the non-budding stage (Equation 1), the right hand side is composed by a negative budding term, representing cells leaving the non-budding stage by initiating the budding process, and a positive birth term, describing the cells entering this stage as a result of the division of a budding cell into two non-budding cells. In the case of the budding stage (Equation 2), the right hand side is composed of the negative division term (cells leaving the stage as result of division into two new cells) and the positive budding term (upon initiation of the budding process, cells transit from the non-budding to the budding stage).

$$\begin{aligned}&\frac{\partial N^{NB}(m,t)}{\partial t} + \frac{\partial}{\partial m}\left[r_m(m,Z)N^{NB}(m,t)\right]\\ &\quad = -\Gamma_B(m|Z)N^{NB}(m,t)\\ &\quad + 2\int_{m}^{m_f}\Gamma_D(m'|Z)P(m,m'|Z)N^{B}(m',t)\mathrm{d}m'\end{aligned} \tag{1}$$

$$\begin{aligned}&\frac{\partial N^{B}(m,t)}{\partial t} + \frac{\partial}{\partial m}\left[r_m(m,Z)N^{B}(m,t)\right]\\ &\quad = -\Gamma_D(m|Z)N^{B}(m,t) + \Gamma_B(m|Z)N^{NB}(m,t)\end{aligned} \tag{2}$$

$$\begin{aligned}N^{NB}(m_0,t) &= N^{NB}(m_f,t) = N^{B}(m_0,t)\\ &= N^{B}(m_f,t) = 0; \quad m\in[m_0,m_f]\end{aligned} \tag{3}$$

$$N^{NB}(m,t=0) = \frac{(1/\sigma_{N^{NB}(m,t=0)})\phi((m-\mu_{N^{NB}(m,t=0)})/\sigma_{N^{NB}(m,t=0)})}{\Phi((m_f-\mu_{N^{NB}(m,t=0)})/\sigma_{N^{NB}(m,t=0)}) - \Phi((m_0-\mu_{N^{NB}(m,t=0)})/\sigma_{N^{NB}(m,t=0)})} \tag{4}$$

$$N^{B}(m,t=0) = \frac{(1/\sigma_{N^{B}(m,t=0)})\phi((m-\mu_{N^{B}(m,t=0)})/\sigma_{N^{B}(m,t=0)})}{\Phi((m_f-\mu_{N^{B}(m,t=0)})/\sigma_{N^{B}(m,t=0)}) - \Phi((m_0-\mu_{N^{B}(m,t=0)})/\sigma_{N^{B}(m,t=0)})} \tag{5}$$

The boundary condition (Equation 3) assumes that the minimum and maximum cell size considered (m_0 and m_f) are sufficiently small and big, respectively, so that the number of cells presenting these sizes is zero. In this work, m_0 is defined as 0 ch. nr. and m_f as 2,000 ch. no. Although the experimental data were collected for ch. no. up to 10^4 (see Supplementary Information S3), the number of cells observed for channel numbers higher than 2,000 was found insignificant, and the discretization upper boundary was therefore set to 2,000 ch. no.

The initial distribution for each of the stages is a necessary condition for the PBEs to be solved (Equation 4). To minimize the influence of how the inoculum was prepared on the measured distributions, and allow for as many generations as possible in the well-controlled environment in the bioreactor, the cell concentration was very low at the beginning of the batch (OD ≈ 0.001). Such low concentrations make the determination of the initial distribution prone to error, due to small sample volumes. The distributions measured for the inocula were not considered, as the growth conditions are substantially different from the ones in the bioreactor, upon inoculation. The initial distribution was thus not measured, but rather assumed to follow a truncated Gaussian distribution with means equal to $\mu_{N^{NB}(m,t=0)} = 500$ and $\mu_{N^{B}(m,t=0)} = 650$ ch. no. for the non-budding and budding stages, respectively, and a standard deviation $(\sigma_{N^{NB}(m,t=0)}, \sigma_{N^{B}(m,t=0)})$ of 100 ch. no. for both stages. Different mean values were tested without yielding significant impact on the model predictions (data not shown). An increase in the standard deviation has an effect on the model predictions for the initial time points (see the Results and Discussion Section).

Growth kernel

The growth rate function, $r_m(m, Z)$ was defined as the product of a mass dependent factor m and a substrate dependent factor $\lambda(Z)$ (Equation 6). First order kinetics were assumed for the mass dependent factor (Hatzis and Porro, 2006; Mantzaris et al., 1999), while the substrate dependent factor $\lambda(Z)$, which can be regarded as a specific growth rate, was derived from the unstructured model for the extracellular environment. The constant k_m operates as a switch that allows for modulating the growth rate in order to reflect the residual growth observed in the stationary phase. For the other cultivation phases, k_m is equal to unity.

$$r_m(m, Z) = k_m m \cdot \lambda(Z) \tag{6}$$

Budding and division kernels

The budding and division rates, Γ_B and Γ_D (Equations 7 and 8), were defined as the product of the growth rate and a hazard function (Hatzis and Porro, 2006; Mantzaris et al., 1999). The latter is based on a density function (h_B or h_D) that describes the probability of a cell of size m to initiate the budding process or to undergo division. Truncated Gaussian probability density functions with mean μ_B or μ_D (the critical transition sizes), respectively, were used. Equation (8) shows the budding probability density function h_i, where φ is a Gaussian probability density function and Φ is a Gaussian cumulative density function. The two mean parameters (μ_B and μ_D) are function of the substrate availability, while the standard deviations (σ_B and σ_D) were assumed to be constant. The numerical values for these parameters are provided as Supplementary Information (Table S4-2).

$$\Gamma_i = r_m(m, Z) \frac{h_i(m|Z)}{1 - \int_{m_0}^{m} h_i(m_k|Z)\mathrm{d}m_k} \quad i = B, D \tag{7}$$

$$h_i(m|Z) = \frac{(1/\sigma_i)\phi((m - \mu_i(Z))/\sigma_i)}{\Phi((m_f - \mu_i(Z))/\sigma_i) - \Phi((m_0 - \mu_i(Z))/\sigma_i)} \quad i = B, D \tag{8}$$

Experimental observations (see the Results and Discussion Section) indicated that during late growth on glucose the mean cell size decreased monotonously without a significant variation of the fraction of budding cells in the population (i.e., BI). The smooth shift of the distribution towards smaller sizes was triggered when a given glucose consumption rate is achieved (see the Results and Discussion Section), and modeled by a linear decrease of the critical transition sizes μ_B and μ_D. The decrease in the critical sizes during the late glucose growth phase is described by Equation (9), where I is a switch equal to 1 when the glucose uptake rate is higher than $0.6\,g\,g^{-1}\,L^{-1}$, and 0 otherwise. A similar behavior was observed during late growth on ethanol (see the Results and Discussion Section), but in this case, the threshold for ethanol uptake rate is $0.15\,g\,g^{-1}\,L^{-1}$. The values of rates k_B and k_D (negative values) of both late growth phases are provided as Supplementary Information (Table S4-2).

$$\frac{\partial \mu_i}{\partial t} = k_i I\left(\frac{\mathrm{d}G}{\mathrm{d}t} < -0.6\right) \quad i = B, D \tag{9}$$

At the end of exponential growth on glucose, the transition from the non-budding to the budding stage is arrested when glucose is depleted, and the diauxic shift is initiated. This budding transition arrest was modeled by imposing a budding rate equal to zero. The duration of this arrest is however not dependent on the model variables, and as such to be defined according to the experimental observations: an arrest of 3 h was defined (reflecting the experimental optical density curve) in the work here presented.

Birth kernel

The birth term describes the formation of two cells of mass m and $m'-m$ as a result of the division of a budding cell of size m'. The birth rate is defined based on the division rate of the mother cell, $\Gamma_D(m'|Z)$, and the partitioning function $P(m, m'|Z)$. The latter function describes the ratio of sizes of the daughter cells (originated from the bud) to the mother budding cell, and is defined as a beta probability density function (Hatzis et al., 1995). In this work, the distribution was set as symmetrical ($\alpha = \beta = 50$) during exponential growth on glucose, while during exponential growth on ethanol it was defined as left-skewed ($\alpha = 30$; $\beta = 60$). The change in the shape parameters reflects the decrease in the ratio of daughter to mother cell size that has been observed experimentally after the diauxic shift (Alberghina and Westerhoff, 2005). The sensitivity of the model output to the beta distribution shape parameters (α and β) defining the partitioning function is discussed further in this work (cf. the Results and Discussion Section).

Unstructured Kinetic Model for the Extracellular Environment

An unstructured kinetic model was used to describe how, during a batch cultivation, glucose (carbon source) is consumed, and how ethanol is first produced, and subsequently consumed after glucose is depleted. Also the consumption of dissolved oxygen was modeled as budding yeast oxidizes or reduces glucose depending on the available dissolved oxygen concentration, and its respiratory capacity. In order to be able to capture the interplay between the cells and composition of the extracellular cultivation environment, a simple kinetic (unstructured) model reported in the literature (Sonnleitner and Käppeli, 1986) was used for describing the respiratory and fermentative growth of budding yeast. In the original model, the biomass accumulation is defined as an autocatalytic reaction. In this work, the corresponding specific growth rate is defined as $\lambda(Z)$, the extracellular environment factor included in the single cell growth rate (Equation 6) in the PBE. Additionally, the biomass concentration is here predicted using the PBM, and provided as input for the prediction of the extracellular environment variables: glucose, ethanol, and oxygen (Fig. 2). The overall biomass concentration at any given time point was calculated based on the sum of the zeroth moment (total number of cells) of the cell size distribution for the two stages. The total number of cells is converted into cell dry weight concentration using a linear regression of cell number to dry weight determined experimentally (data not shown). Although it would be expected that cells of small size present a lower dry weight than bigger cells, the error typically associated with the experimental dry weight determination may be quite significant and it would shade differences between smaller and bigger cells. Therefore, the total number of cells was converted into dry weight concentration using a linear regression of cell number to cell dry weight determined experimentally (data not shown).

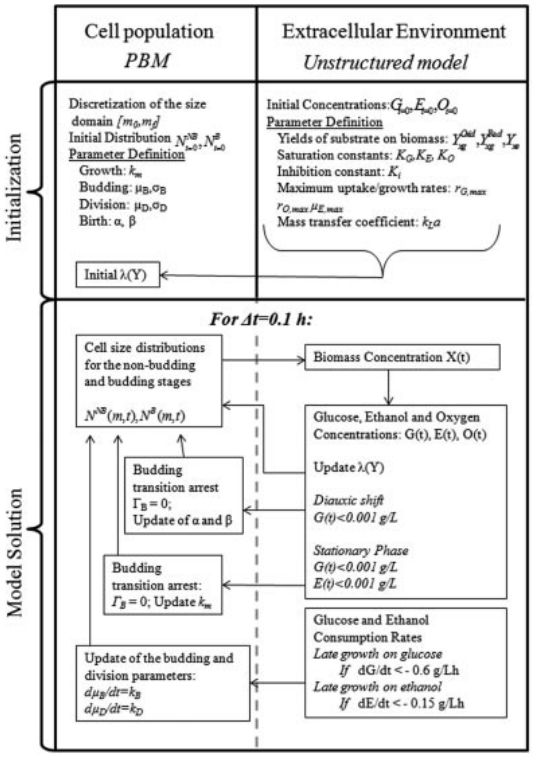

Figure 2. Schematic representation of the numerical procedure for solution of the PBM and the coupled unstructured kinetic model.

The average yield coefficients were estimated by fitting of the model to the experimental data, and are provided as Supplementary Information (Table S4-1) together with a description and the equations for this unstructured model (Supplementary Information S5). Although ideally desirable, the parameter estimation was not performed by an optimization routine due to the high collinearity of the parameters in the unstructured model and the complexity of the overall multi-scale model. The design of a parameter estimation routine using a multivariate objective function would be a valuable contribution, it however falls out of the scope of this work.

Model Solution Methods

The fixed-pivot technique (Kumar and Ramkrishna, 1996; Nopens et al., 2005) was applied to an evenly distributed discretization grid with 166 pivots (in the range $[m_0, m_f] = [0, 2000]$). The PBM and the unstructured kinetic model are solved iteratively for time steps of 0.1 h, where the biomass concentration estimated by the PBM is supplied to the unstructured model, and the updated concentrations of

glucose, ethanol and oxygen are used to re-calculate the substrate dependent factor $\lambda(Z)$ in the growth rate, and eventually update the budding and division parameters μ_B and μ_D, which are depending on the substrate consumption rates. A schematic representation of the model solution steps is presented in Figure 2.

The time step size was defined in order to ensure a variation in the concentrations of glucose and ethanol per time step smaller than $0.1\,g\,L^{-1}$—the approximate experimental error in the determination of these concentrations (see the Materials and Methods Section). Smaller time steps did not yield significant differences in the model predictions (data not shown), confirming solution convergence. All model simulations were performed using MatLab® R2009b (The MathWorks, Inc.).

Results and Discussion

Triplicate glucose-limited aerobic batch cultivations of *S. cerevisiae* were performed. The three cultivations presented similar growth and production profiles (Fig. S6-1 in Supplementary Information). The respective biomass (optical density), glucose and ethanol concentration curves followed the typical patterns for aerobic yeast cultivations. The population proliferated exponentially first on glucose with a specific growth rate of $0.41 \pm 0.001\,h^{-1}$, and then on ethanol with a growth rate of $0.10 \pm 0.02\,h^{-1}$. The distributions of cell total protein content (a measure of cell size) and DNA were monitored during the four cultivation phases by flow cytometry. Generally, there was a good agreement between the three replicate cultivations with regards to the flow cytometry measurements (Fig. S6-2 in Supplementary Information).

A model consisting of both a population balance model (PMB) part and an unstructured kinetic model part was constructed (for details see the section on Model Aspects above). The predictions of this integrated model describe both the variation of macroscopic variables—glucose, ethanol, and dissolved oxygen, as well as the total biomass concentration—and the distributions of single-cell total protein content for the non-budding and budding populations during glucose-limited batch cultivation. With regard to the macroscopic variables, the comparison of the model predictions and experimental results (Fig. 3) shows a good agreement: the model is able to describe the growth phases (on glucose and ethanol) as well as the intermediate diauxic shift phase. Based on the increase of number of cells predicted by the model, the model predicted specific growth rate was $0.44\,h^{-1}$. This corresponds to a deviation of approximately 7% between the estimation and experimental data.

Mean Total Protein Content

The mean total protein content of the population decreased throughout the cultivation (Fig. 4). During exponential growth on glucose, the population showed a higher mean total protein content relative to the other cultivation phases (i.e., larger cell size), corresponding to a channel number of approximately 700. In late exponential phase during glucose assimilation, the mean cell size decreased to a ch. no. of approximately 450. An additional decrease of the mean cell size to a ch. no. of 350 was observed upon the depletion of glucose. Similarly to what was observed during growth on

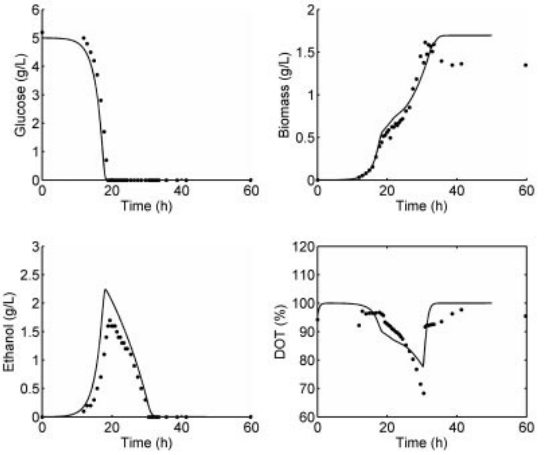

Figure 3. Variation of glucose, total biomass, and ethanol concentrations and dissolved oxygen tension along the cultivation: model predictions (full line) and experimental observations (full circles).

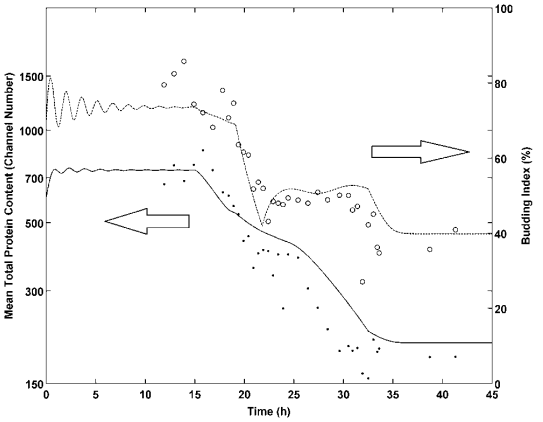

Figure 4. Comparison of model predictions for the mean cell size or total protein content (full line) and budding index (dashed line) to experimental observations (cell size as dots and budding index as open circles).

glucose, the mean cell size decreased during late exponential growth on ethanol, from a mean ch. no. of approximately 300–200. The mean cell size remained approximately constant during stationary phase until $t = 60$ h, when the cultivations were terminated.

Furthermore, also the predicted mean total protein content for the overall population was in good agreement with the experimental observations.

Budding Index

The variation in the distribution of cells through the cell cycle phases can be seen even clearer when observing the variation of the BI or the fraction of budding cells in the population, along the cultivation (Fig. 4).

Estimation of BI from microscopic images was performed to verify the estimated BI from flow cytometry DNA distributions. Even if there were deviations in the exact numbers, a similar trend was observed (see Supplementary Information S2). Moreover, the BIs estimated from flow cytometric data for time points during exponential growth on glucose and ethanol, presented in this work, are in good agreement with values reported in the literature. A BI of approximately 80% was observed during growth on glucose, which is in agreement with reported BI values for cultivations with similar specific growth rates (Brauer et al., 2005; Cipollina et al., 2005).

The initial oscillations of the BI and subsequent damping can be interpreted as a desynchronization of the population resulting from the fact that the critical sizes are described by a probability distribution function rather than a unique discrete value (Liou et al., 1997; Mantzaris et al., 1999). An increase in the standard deviation assumed for the initial distribution results in a quicker damping of these oscillations.

The abrupt decrease of the BI to approximately 48% from about $t = 18$ h onwards coincides with the depletion of glucose. A similar BI has been observed during exponential growth in an ethanol-limited batch cultivation (Cipollina et al., 2005). During stationary phase, an accumulation of G1 cells took place analogously to the observations during the diauxic shift, resulting in a BI of approximately 42% after 60 h of cultivation.

Generally, the model predictions are in good agreement with the estimations based on the experimental DNA distributions here reported. Indeed, the accumulation of smaller non-budding cells observed in the diauxic shift is successfully described by the model, and reflected in the continued decrease of the mean cell size and accentuated the decrease of the BI during this phase of metabolic rearrangement (Fig. 4).

When ethanol is depleted and cells enter the stationary phase, the typically observed accumulation of non-budding cells takes place at slower pace than the one observed for the diauxic shift. Therefore, an arrest of the budding transition is imposed as well as decreasing the constant k_m from unity to an assumed value of 0.4. Notwithstanding the capability of the model of describing the decrease in the mean cell size and BI, the transition into stationary phase is rather complex, and not yet fully understood (Werner-Washburne et al., 2011). Therefore, the simplistic model description presented (by decreasing the k_m and arresting the budding transition) should not be expected to describe such complexity. More research is required to fully unravel this mechanism.

Cell Size and Cell Cycle Position Distributions

In Figure 5, the bivariate distributions of total protein and DNA content measured along the cultivation are presented as contour plots. The distributions of each cell property (in the shape of histograms), for all sample times, are provided as Supplementary Information (S7).

From the bivariate distribution, the significant changes in the structure of the population can be easily visualized. During the growth on glucose, a large part of the population consisted of bigger cells presenting two copies of the DNA (G2 + M). As reflected in the BI profile (Fig. 4), during the diauxic shift, an accumulation of cells with lower DNA content (G1) and smaller size was observed, resulting in a clearer cloud on the bottom left corner of the bivariate distribution (Fig. 5, $t = 20.4$ h). During growth on ethanol, a more even distribution of the cells containing one or two copies of DNA, in comparison to growth on glucose, was observed. This is illustrated in Figure 5: while for exponential growth on glucose ($t = 16.8$ h) only one high density cloud (red colored) is observed, two clouds with approximately the same densities (blue colored) are observed during the growth on ethanol ($t = 23.9$ h).

In Figure 6, the total protein content distributions predicted by the model are compared to the experimental ones. The model predictions successfully describe the general shift of the distributions towards smaller sizes. The largest difference between model predictions and experimental distributions is observed during the diauxic shift. While the model shows clear bimodal distributions around $t = 21$ h, the experimental distributions are unimodal. The experimental distributions for each of the non-budding and budding subpopulations may present larger variances than described in the model, and this may explain the loss in separation of the peaks corresponding to non-budding and budding cells (unimodality) observed experimentally. Additionally, the model does not acknowledge the existence of different generations whose critical sizes increase with the generation age (Hatzis and Porro, 2006), leading to intermediate subpopulations.

Critical Budding and Division Sizes: Dependence on Substrate Availability and Uptake

The model presented in this work relies on the fact that the cell size regulation has been identified as a key aspect in

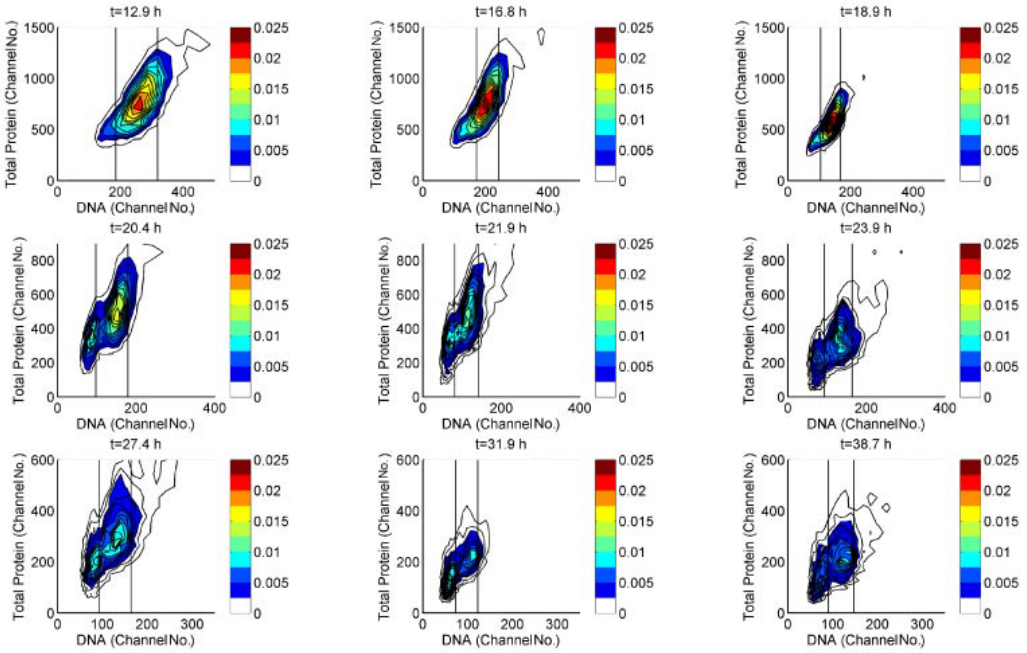

Figure 5. Bivariate distribution of total protein content and DNA during a batch cultivation of *S. cerevisiae*: exponential growth on glucose ($t = 12.9$, 16.8 h), diauxic shift ($t = 18.9$, 20.4 h), exponential growth on ethanol ($t = 21.9$, 23.9, 27.4 h), and stationary state ($t = 31.9$, 38.7 h). The color code corresponds to the number density of the cells. The vertical lines correspond to the critical budding (to the left) and division (to the right) threshold identified based on the DNA distribution.

the growth regulation in response to changing environmental conditions (Rupeš, 2002; Vanoni et al., 2005). Understanding the dependence of the critical budding and division sizes on the extracellular environment is essential for understanding the population dynamics and for future modeling of the dynamics of size and cell cycle position distributions. Indeed, to define the critical budding and division sizes as a function of the substrate availability is essential for the definition of a dynamic population model (cf. Modeling Aspects Section).

In order to assess how the mean cell size of a subpopulation with a high content of cells initiating the budding process, that is, cells in the critical budding size range, changed during the batch cultivation a standardized procedure was developed and applied to all samples (see Computational Data Treatment Section). Not surprisingly the critical budding size followed a similar evolution as the overall mean cell size (Fig. 7), and was constant during early growth phases and decreased in the later growth phases and diauxic shift. Similarly, a standardized procedure was applied to estimate the critical division size at each sample time point. Also the critical division size accompanied the general shift of the overall population towards smaller sizes along the cultivation (Fig. 7).

In this work, particular attention was paid to understand the population dynamics observed during the transition between growth on glucose and ethanol. From the variation of the estimated critical budding and division sizes along the cultivation during the late growth phases on glucose or ethanol (Fig. 7), cells seem to adjust to the decreasing substrate availability by a smooth shift of the cell size distributions towards smaller sizes, and a very slight decrease in the BI (Fig. 4). Contrarily, an abrupt change in the cell cycle position (sudden sharp decrease of the BI) was observed upon glucose depletion and beginning of the diauxic shift. These observations are in agreement with the two different mechanisms proposed by Brauer et al. (2005) when explaining the changes in the gene expression patterns observed during (i) the late growth phases—"continuous homeostatic metabolic adjustment"—and (ii) diauxic shift—"discontinuous metabolic remodeling."

Cell Adjustment During Late Growth Phases

As mentioned above, during the late growth phases, the cell size distribution smoothly shifted towards smaller sizes, while the BI decreased very slightly (Fig. 4). In this work, the initial glucose concentration was 5 g L^{-1}, and the decrease in

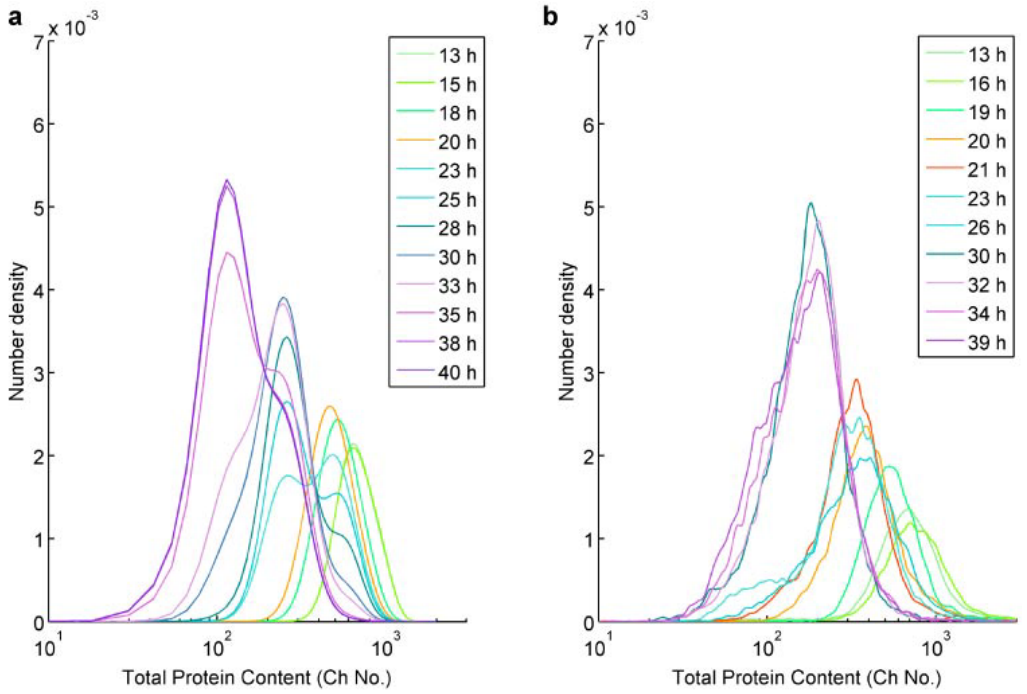

Figure 6. Total protein content distributions for the overall population: (**a**) model predictions (**b**) experimental observations. The color code reflects the different cultivation phases: exponential growth on glucose (green), diauxic shift (orange-red), exponential growth on glucose (blue), stationary phase (violet).

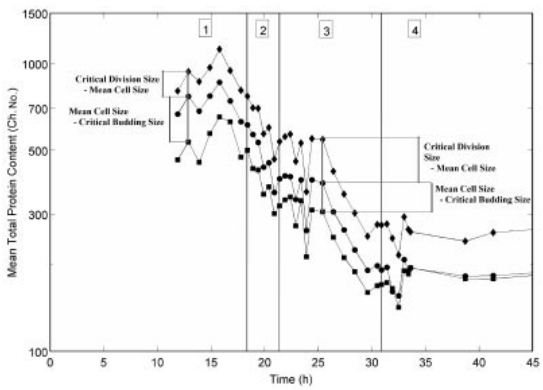

Figure 7. Variation of the mean total protein content (i.e., cell size) for the overall population (circles), the critical budding size (squares), and critical division size (diamonds). The vertical lines define the four phases of the cultivation: (1) exponential growth on glucose; (2) diauxic shift; (3) exponential growth on ethanol; (4) stationary phase. The distances between cell size and the critical budding and division sizes reflect the cell cycle distribution (fractions of non-budding and budding cells within the population).

cell size was observed approximately 2 h before the diauxic shift occurred, corresponding to a glucose concentration above 3 g L^{-1}. Brauer et al. (2005) performed cultivations with an initial glucose concentration of 2.5 g L^{-1}, and significant changes in gene expression were observed when the glucose concentration was approximately 0.5 g L^{-1}. It thus seems unlikely that the onset of the metabolic adjustment is triggered by a certain concentration of glucose in the medium.

Interestingly, when comparing the glucose consumption and biomass growth curves of Brauer et al. (2005) with those presented in Figure 3, the start of the adjustment in gene expression and decrease in the mean cell size coincided with a steeper increase in biomass concentration (OD) and consumption of glucose. The glucose consumption rate estimated for this time point was approximately 0.6 g L^{-1} h^{-1}. A similar rate was estimated from the publication of Brauer et al. (2005). With regard to the growth phase on ethanol, the trend of cell size distribution and the critical budding size was similar to the behavior during growth on glucose. A similar adjustment behavior triggered by threshold ethanol consumption rate was thus assumed. This threshold was estimated to be approximately 0.15 g L^{-1} h^{-1}. These thresholds were used in the model for

defining dependence of the critical budding and division sizes as functions of the substrate uptake (see section on the Modeling Aspects and Fig. 2).

The dependence of the critical budding size on the specific growth rate has been reported elsewhere (Porro et al., 2003), and is implicitly related with the uptake rates (Youk and van Oudenaarden, 2009). The link between the critical division size and the uptake rate is thus not unexpected. In fact, Youk and van Oudenaarden (2009) discuss the interaction between glucose perception and import (uptake) on the growth: for wild-type *S. cerevisiae*, when glucose is not too low, the effect of glucose perception disappears and the uptake perception is dominant. A more in depth validation of these assumptions and numerical values for the threshold uptake rates here proposed could, for example, be achieved by comparison to the behavior of strains with improved glucose or ethanol affinities. Such validation lies, however, outside the scope of this work.

Cell Rearrangement Upon Diauxic Shift

During the diauxic shift, an abrupt decrease of BI was observed (Fig. 4) together with the appearance of a subpopulation of smaller cells (Fig. 5). The cell rearrangement is likely to have more underlying reasons than merely a modulation of the critical budding and division sizes. In fact, it has been reported that the difference in cell size between the two cells originating upon division—the bigger mother cell and the smaller new born cell—is more accentuated for growth on ethanol than for growth on glucose (Di Talia et al., 2007, 2009). Furthermore, the diauxic shift is typically characterized by arrest in the increase of biomass. As reviewed by Alberghina and Westerhoff (2005), the budding transition is likely to be resumed before the division, leading to a small increase of the BI after a sharp downward shift. This can be described by an abrupt reset of the partition parameters, and an arrest of the budding transition. A small downward peak and consequent recovery were observed in Figure 4 ($t = 22.9$ h); however it should not be considered as conclusive given the nature of the BI estimation procedure. Such an abrupt change in the partition shape parameters allowing smaller cells to eventually initiate DNA replication and the budding process were however not captured when considering the mean critical budding size (Fig. 7). Indeed, the standardized procedure developed in this work based on analyzing the DNA histograms is not able to capture abrupt dynamics: it is necessary that a sufficient fraction of cells with a smaller size initiating DNA replication is present in the population, for its effect on the mean size of the subpopulation isolated as the critical budding band to be observed. The sensitivity of the method may be improved by increasing the number of analyzed cells for the cases when faster dynamics are expected. Furthermore, cells of different ages could be distinguished experimentally by additional staining (Sumner et al., 2003), allowing specifically for isolation and comparison of the newborn cells originated in the different growth phases.

Sensitivity of the Model Output to the Partition Function Parameters

The description of the partition function as a beta distribution implies that the two shape parameters, α and β, are defined. In a previous theoretical study by Mantzaris et al. (2002), a symmetrical distribution with shape parameters, α and β, equal to 40 was assumed without experimental evidences. Other studies (Fadda et al., 2012b; Hatzis and Porro, 2006; Mancuso et al., 2009; Sidoli et al., 2006) assumed the same symmetrical distribution and shape parameter.

In this study, the numerical values of each of the shape parameters have also not been directly estimated from the experimental data, but the effect of these parameters on the model outputs has been assessed. An experimental estimation of the partition distribution (and corresponding shape parameters) implies determining the cell size distribution of a subpopulation of newly born cells: not only the smaller cells originating from the bud but also the bigger cells that produced the bud. This could be done taking a similar approach to Porro and Srienc (1995): at a given time point all cells in a sample were stained with ConA-FITC (labeling the cell surface) and resuspended in growth media. Assuming that the first partly stained cells to appear correspond to cells that have just divided, the total protein content of this subpopulation can be determined: a symmetric distribution will indicate a symmetric partitioning (more equally sized cells) while an asymmetric (eventually bimodal) distribution would indicate a unequal partitioning. It is however not clear if this experimental procedure would allow for a quantitative estimation of the shape parameters.

The numerical values chosen—a symmetrical distribution initially ($\alpha = \beta = 50$) and a left-skewed distribution from the diauxic shift onwards ($\alpha = 30$ and $\beta = 60$)—correspond to the best fit of the model predictions to the experimental data. The change in the parameters from one growth phase to the next, reflects the fact that the two cells originated upon division are more similar in size during growth on glucose than during growth on ethanol.

Model simulations for different parameter combinations were compared in order to assess the sensitivity of the model to the two shape parameters. Firstly, simulations with different combinations of values for α and β during growth of glucose were carried out, while maintaining those after the diauxic shift at $\alpha = 30$ and $\beta = 60$ (Case I). The resulting model predictions are compared in Figure 8. Secondly, simulations with different combinations of values for α and β after the diauxic shift were made, while maintaining the initial values $\alpha = \beta = 50$ for growth on glucose constant (Case II). The resulting model predictions are compared in Figure 9. The beta distributions corresponding to the various parameter combinations are illustrated as Supplementary Information (Figs. S8-1 and S8-2).

Concerning the sensitivity of the model to the partitioning function parameters provided initially (Case I, Fig. 8),

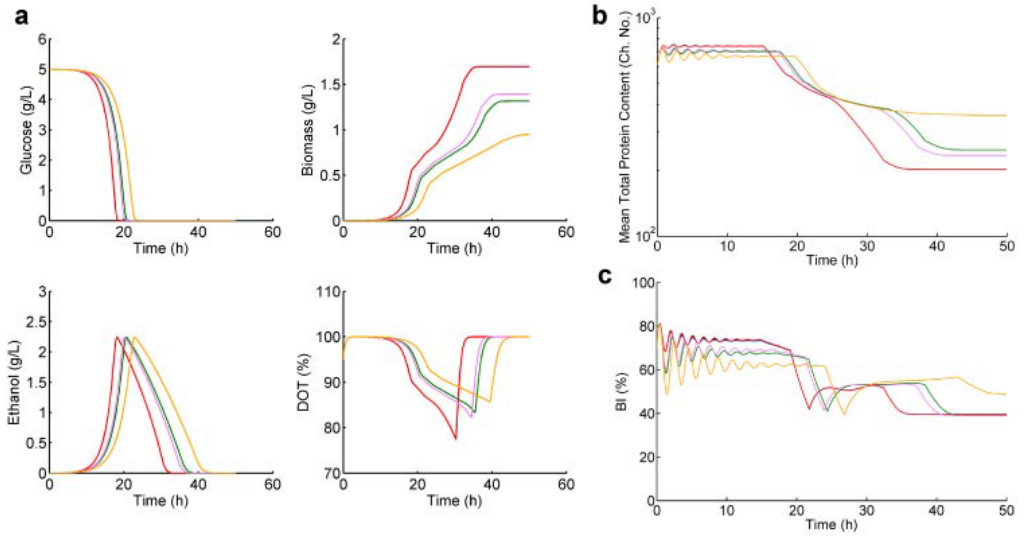

Figure 8. Comparison of the model outputs for different combinations of the shape parameters in the partitioning function: $\alpha = \beta = 50$ (black), $\alpha = \beta = 40$ (blue), $\alpha = \beta = 60$ (red), $\alpha = 40$ and $\beta = 50$ (green), $\alpha = 50$ and $\beta = 60$ (pink), and $\alpha = 40$ and $\beta = 60$ (orange). **a**: Variation of the macroscopic variables: glucose, ethanol, biomass, and oxygen; (**b**) variation of the mean total protein content; (**c**) variation of the budding index (BI). The black and the blue lines coincide with the red one.

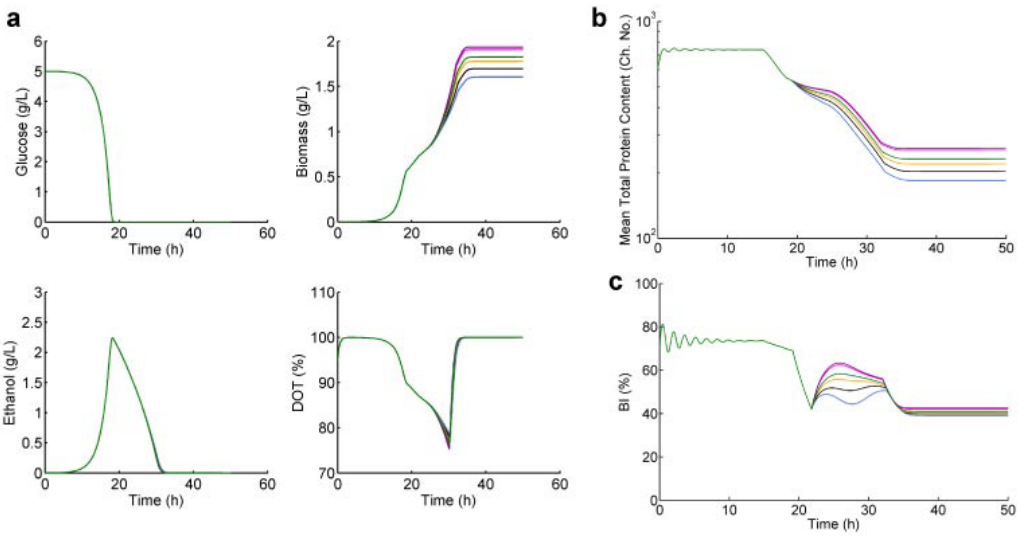

Figure 9. Comparison of the model outputs for different combinations of the shape parameters of the partitioning function as of the diauxic shift: $\alpha = 30$ and $\beta = 60$ (black), $\alpha = 20$ and $\beta = 50$ (blue), $\alpha = 30$ and $\beta = 50$ (gold), $\alpha = 40$ and $\beta = 50$ (magenta), $\alpha = 50$ and $\beta = 60$ (purple), $\alpha = 40$ and $\beta = 60$ (green). **a**: Variation of the macroscopic variables: glucose, ethanol, biomass, and oxygen; (**b**) variation of the mean total protein content; (**c**) variation of the budding index (BI). The lines in magenta and purple nearly coincide.

the more asymmetrical the beta distribution (i.e., bigger difference in the size of the two cells resulting from division) the slower the growth is. The slower growth leads to a slower consumption of glucose and a consequent time delay in the ethanol and dissolved oxygen profiles (Fig. 8a). This can be explained by the existence of smaller cells (lower overall mean cell size—Fig. 8b) which require a longer time to go through the cell cycle, that is, spend more time in the non-budding stage (lower BI—Fig. 8c). Furthermore, the existence of smaller non-budding cells results in greater oscillations of the BI (Fig. 8c). As expected, a similar effect is observed when different parameter combinations are used for redefining the partitioning function upon the diauxic shift (Case II). The sensitivity of the ethanol and dissolved oxygen concentrations (model outputs), respectively, are however lower in this case (Fig. 9a). After the abrupt decrease caused by the arrest in the budding transition, the BI increases upon resuming the transition: the more symmetrical the partition function, the higher the predicted BI is upon resuming the budding transition. This is not surprising when considering that, for a more symmetrical distribution, the cells accumulating in the non-budding stage (during the arrest of the budding transition) have more similar sizes and, upon resuming the budding transition, will become budding cells at approximately the same time causing a larger increase in the BI.

Conclusions

This study focuses on understanding and modeling the dynamics of a yeast population in terms of development of the cell size and cell cycle distributions along a batch cultivation. It presents of an example of a quantitative integrated analysis of general physiology data (i.e., substrate, metabolite and biomass concentrations) and single-cell flow cytometry data. The experimental single-cell measurements for DNA and total protein, as well as the concentrations of substrate and metabolites were interpreted also in the light of gene expression studies previously reported in the literature.

The standardized procedures developed allow for identifying trends in the single-cell properties (critical sizes) along the cultivation. Although, these procedures require the definition of a threshold (e.g., distance from the peak mode) and this decision may introduce a bias, the same threshold decision is made for all the samples, implying that the bias is systematic. Oppositely, manual gating for each sample would introduce a random bias, while the use of the same fixed gates for all samples would neglect the fact that the overall distribution shifts along time also introducing a varying bias. Hence, standardized procedures such as those developed in this work are important for quantitative, systematic interpretation of data from single-cell analysis and for use of this type of data in mathematical models.

Population balance models offer a framework to describe the dynamics of distributed properties in a wide range of applications. Although the first discussion on the use of such models for describing microbial populations was published more than 40 years ago (Fredrickson et al., 1967), reports on the application of the models to specific examples and comparison to experimental results are scarce (Lencastre Fernandes et al., 2011). One of the causes for such a limited use of these segregated models, in a time where experimental methods for measuring single-cell properties (e.g., flow cytometry, fluorescence microscopy) are easily available, is the difficulty in translating observed behaviors into the PBM growth, budding, division and birth terms.

The overall data analysis allowed for formulating two key kernel functions—budding and division rates—based on dependence of the critical budding and division sizes on the glucose and ethanol uptake rates.

Furthermore, the coupling of the PBM to an unstructured kinetic model, proposed in this work, results in a more comprehensive description of the phenomena taking place at different scales (macroscale and microscale) during the cultivation. Such coupling is essential for integrating the typical physiological data (averaged measurements for the overall population), with distributed data collected at single-cell level, thus achieving a model with improved prediction capacities at two levels of detail.

The presented work contributes towards linking experimental data and PBM theoretical work: a new trend in the PBM community (Nopens et al., 2009). The proposed model can be regarded as a tool for investigating these dynamics under different scenarios, for example, pulse experiments, and comparing the different assumptions against the experimental observations. In the case of continuous or fed-batch fermentations where the glucose concentration in the feed is high, the integration of the proposed model and a computational fluid dynamics (CFD) model describing the distribution of substrate within the reactor provides a valuable tool to study in silico the effect of non-ideal mixing, and resulting substrate gradients, on the development of heterogeneous cell populations. Recently, the occurrence of distribution of protein levels (concentrations) for a cell population was studied in silico using a PBM that incorporated a stochastic description for gene expression (Shu et al., 2011, 2012). It was concluded that, although it is often believed that the occurrence of bimodal distributions results from a bistability of the gene regulatory network (e.g., Hasty et al., 2002; Thattai and van Oudenaarden, 2001), this bimodality may arise even if the stochastic bistability does not occur (Shu et al., 2011). This interesting observation indicates that bimodality may be a consequence of the dynamics of a population. A PBM including a deterministic, rather than a stochastic description of the protein production kinetics may thus eventually be sufficient to describe the distribution of protein levels in a microbial population. Moreover, as the interplay with the extracellular environment (e.g., available substrate) could be included, such a model would be a valuable contribution towards understanding the impact of the development of heterogeneous populations on the overall productivity, and

obtain a better predictive tool for modeling large-scale fermentors where heterogeneity has been observed (Enfors et al., 2001) and still is a pending problem in practice.

References

Alberghina L, Westerhoff HV, editors. 2005. Systems biology, Vol. 13. Berlin/Heidelberg: Springer-Verlag.

Alberghina L, Smeraldi C, Ranzi BM, Porro D. 1998. Control by nutrients of growth and cell cycle progression in budding yeast, analyzed by double-tag flow cytometry. J Bacteriol 180:3864–3872.

Avery SV. 2006. Microbial cell individuality and the underlying sources of heterogeneity. Nat Rev Micro 4:577–587.

Brauer MJ, Saldanha AJ, Dolinski K, Botstein D. 2005. Homeostatic adjustment and metabolic remodeling in glucose-limited yeast cultures. Mol Biol Cell 16:2503–2517.

Carlquist M, Lencastre Fernandes R, Helmark S, Heins A-L, Lundin L, Sørensen SJ, Gernaey KV, Eliasson Lantz A. 2012. Physiological heterogeneities in microbial populations and implications for physical stress tolerance. Microb Cell Fact 11:94.

Cipollina C, Alberghina L, Porro D, Vai M. 2005. Sfp1 is involved in cell size modulation in respiro-fermentative growth conditions. Yeast 22: 385–399.

Cipollina C, Vai M, Porro D, Hatzis C. 2007. Towards understanding of the complex structure of growing yeast populations. J Biotechnol 128:393–402.

Davey H, Kell D. 1996. Flow cytometry and cell sorting of heterogeneous microbial populations: The importance of single-cell analyses. Microbiol Rev 60:641–696.

Di Talia S, Skotheim JM, Bean JM, Siggia ED, Cross FR. 2007. The effects of molecular noise and size control on variability in the budding yeast cell cycle. Nature 448:947–951.

Di Talia S, Wang H, Skotheim JM, Rosebrock AP, Futcher B, Cross FR. 2009. Daughter-specific transcription factors regulate cell size control in budding yeast. PLoS Biol 7:e1000221.

Enfors S-O, Jahic M, Rozkov A, Xu B, Hecker M, Jürgen B, Krüger E, Schweder T, Hamer G, O'Beirne D, Noisommit-Rizzi N, Reuss M, Boone L, Hewitt C, McFarlane C, Nienow A, Kovacs T, Trägårdh C, Fuchs L, Revstedt J, Friberg PC, Hjertager B, Blomsten G, Skogman H, Hjort S, Hoeks F, Lin H-Y, Neubauer P, van der Lans R, Luyben K, Vrabel P, Manelius Å. 2001. Physiological responses to mixing in large scale bioreactors. J Biotechnol 85:175–185.

Fadda S, Cincotti A, Cao G. 2012a. A novel population balance model to investigate the kinetics of in vitro cell proliferation: Part I. Model development. Biotech Bioeng 109:772–781.

Fadda S, Cincotti A, Cao G. 2012b. A novel population balance model to investigate the kinetics of in vitro cell proliferation: Part II—Numerical solution, parameters' determination and model outcomes. Biotech Bioeng 109:772–781.

Fredrickson AG. 2003. Population balance equations for cell and microbial cultures revisited. AIChE J 49:1050–1059.

Fredrickson AG, Ramkrishna D, Tsuchiya HM. 1967. Statistics and dynamics of procaryotic cell populations. Math Biosci 1:327–374.

Fredrickson AG, Megee RD, III, Tsuchiya HM. 1970. Mathematical models for fermentation processes, In: Umbreit WW, Perlman D, editors. Advances in applied microbiology, Volume 13. New York: Academic Press. p 419–465.

Hasty J, McMillen D, Collins JJ. 2002. Engineered gene circuits. Nature 420:224–230.

Hatzis C, Porro D. 2006. Morphologically-structured models of growing budding yeast populations. J Biotechnol 124:420–438.

Hatzis C, Srienc F, Fredrickson AG. 1995. Multistaged corpuscular models of microbial growth: Monte Carlo simulations. Biosystems 36:19–35.

Howlett NG, Avery SV. 1999. Flow cytometric investigation of heterogeneous copper-sensitivity in asynchronously grown *Saccharomyces cerevisiae*. FEMS Microbiol Lett 176:379–386.

Jorgensen P, Tyers M. 2004. How cells coordinate growth and division. Curr Biol 14:R1014–R1027.

Jorgensen P, Nishikawa JL, Breitkreutz B-J, Tyers M. 2002. Systematic identification of pathways that couple cell growth and division in yeast. Science 297:395–400.

Kale SP, Jazwinski SM. 1996. Differential response to UV stress and DNA damage during the yeast replicative life span. Dev Genet 18: 154–160.

Karra S, Sager B, Karim MN. 2010. Multi-scale modeling of heterogeneities in mammalian cell culture processes. Ind Eng Chem Res 49:7990–8006.

Kumar S, Ramkrishna D. 1996. On the solution of population balance equations by discretization—i. A fixed pivot technique. Chem Eng Sci 51:1311–1332.

Lencastre Fernandes R, Nierychlo M, Lundin L, Pedersen AE, Puentes Tellez PE, Dutta A, Carlquist M, Bolic A, Schäpper D, Brunetti AC, Helmark S, Heins A-L, Jensen AD, Nopens I, Rottwitt K, Szita N, van Elsas JD, Nielsen PH, Martinussen J, Sørensen SJ, Lantz AE, Gernaey KV. 2011. Experimental methods and modeling techniques for description of cell population heterogeneity. Biotechnol Adv 29:575–599.

Liou J-J, Srienc F, Fredrickson AG. 1997. Solutions of population balance models based on a successive generations approach. Chem Eng Sci 52:1529–1540.

Liu Y-H, Bi J-X, Zeng A-P, Yuan J-Q. 2007. A population balance model describing the cell cycle dynamics of myeloma cell cultivation. Biotechnol Prog 23:1198–1209.

Mancuso L, Liuzzo MI, Fadda S, Pisu M, Cincotti A, Arras M, Desogus E, Piras F, Piga G, La Nasa G, Concas A, Cao G. 2009. Experimental analysis and modelling of in vitro proliferation of mesenchymal stem cells. Cell Prolif 42:602–616.

Mantzaris NV, Liou J-J, Daoutidis P, Srienc F. 1999. Numerical solution of a mass structured cell population balance model in an environment of changing substrate concentration. J Biotechnol 71:157–174.

Mantzaris NV, Daoutidis P, Srienc F. 2001. Numerical solution of multivariable cell population balance models: i. Finite difference methods. Comp Chem Eng 25:1411–1440.

Mantzaris NV, Srienc F, Daoutidis P. 2002. Nonlinear productivity control using a multi-staged cell population balance model. Chem Eng Sci 57:1–14.

Müller S, Harms H, Bley T. 2010. Origin and analysis of microbial population heterogeneity in bioprocesses. Curr Opin Biotechnol 21:100–113.

Nopens I, Beheydt D, Vanrolleghem PA. 2005. Comparison and pitfalls of different discretised solution methods for population balance models: A simulation study. Comp Chem Eng 29:367–377.

Nopens I, Briesen H, Ducoste J. 2009. Celebrating a milestone in population balance modeling. Chem Eng Sci 64:627.

Porro D, Srienc F. 1995. Tracking of individual cell cohorts in asynchronous *Saccharomyces cerevisiae* populations. Biotechnol Prog 11:342–347.

Porro D, Brambilla L, Alberghina L. 2003. Glucose metabolism and cell size in continuous cultures of *Saccharomyces cerevisiae*. FEMS Microb Lett 229:165–171.

Porro D, Vai M, Vanoni M, Alberghina L, Hatzis C. 2009. Analysis and modeling of growing budding yeast populations at the single cell level. Cytometry 75A:114–120.

Rupeš I. 2002. Checking cell size in yeast. Trends Genet 18:479–485.

Saldanha AJ, Brauer MJ, Botstein D. 2004. Nutritional homeostasis in batch and steady-state culture of yeast. Mol Biol Cell 15:4089–4104.

Shu C-C, Chatterjee A, Dunny G, Hu W-S, Ramkrishna D. 2011. Bistability versus bimodal distributions in gene regulatory processes from population balance. PLoS Comput Biol 7:e1002140.

Shu C-C, Chatterjee A, Hu W-S, Ramkrishna D. 2012. Modeling of gene regulatory processes by population-mediated signaling: New applications of population balances. Chem Eng Sci 70:188–199.

Sidoli FR, Asprey SP, Mantalaris A. 2006. A coupled single cell-population model for mammalian cell cultures. Ind Eng Chem Res 45:5801–5811.

Sonnleitner B, Käppeli O. 1986. Growth of *Saccharomyces cerevisiae* is controlled by its limited respiratory capacity: Formulation and verification of a hypothesis. Biotechnol Bioeng 28:927–937.

Sumner ER, Avery AM, Houghton JE, Robins RA, Avery SV. 2003. Cell cycle- and age-dependent activation of sod1p drives the formation of stress resistant cell subpopulations within clonal yeast cultures. Mol Microbiol 50:857–870.

Thattai M, van Oudenaarden A. 2001. Intrinsic noise in gene regulatory networks. Proc Natl Acad Sci 98:8614–8619.

Vanoni M, Vai M, Popolo L, Alberghina L. 1983. Structural heterogeneity in populations of the budding yeast *Saccharomyces cerevisiae*. J Bacteriol 156:1282–1291.

Vanoni M, Rossi RL, Querin L, Zinzalla V, Alberghina L. 2005. Glucose modulation of cell size in yeast. Biochem Soc Trans 33:294–296.

Verduyn C, Postma E, Scheffers WA, Van Dijken JP. 1992. Effect of benzoic acid on metabolic fluxes in yeasts: A continuous-culture study on the regulation of respiration and alcoholic fermentation. Yeast 8:501–517.

Werner-Washburne M, Roy S, Davidson GS. 2011. Aging and the survival of quiescent and non-quiescent cells in yeast stationary-phase cultures. In: Breitenbach M, Jazwinski SM, Laun P, editors. Aging research in yeast, Vol. 57. Dordrecht: Springer Netherlands. p 123–143.

Youk H, van Oudenaarden A. 2009. Growth landscape formed by perception and import of glucose in yeast. Nature 462:875–880.

I.A. Karimi and Rajagopalan Srinivasan (Editors), Proceedings of the 11th International Symposium on Process Systems Engineering, 15-19 July 2012, Singapore.

Multi-scale modeling for prediction of distributed cellular properties in response to substrate spatial gradients in a continuously run microreactor

Rita Lencastre Fernandes[a], Ulrich Krühne[a], Ingmar Nopens[b], Anker D. Jensen[c], Krist V. Gernaey[a]

[a]*Center for Process Engineering and Technology, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark*
[b] *BIOMATH, Department of Mathematical Modelling, Statistics and Bioinformatics, Ghent University, Coupure Links 653, 9000 Ghent, Belgium*
[c] *Center for Combustion and Harmful Emission Control, Department of Chemical and Biochemical Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark*

Abstract

In large-scale fermentors, non-ideal mixing leads to the development of heterogeneous cell populations. This cell-to-cell variability may explain the differences in e.g. yields for large- and lab-scale cultivations. In this work the anaerobic growth of *Saccharomyces cerevisiae* in a continuously run microbioreactor is simulated. A multi-scale model consisting of the coupling of a population balance model, a kinetic model and a flow model was developed in order to predict simultaneously local concentrations of substrate (glucose), product (ethanol) and biomass, as well as the local cell size distributions.

Keywords: Population Balance Model, Computational Fluid Dynamics, yeast, microreactor, fermentation

1. Introduction

A heterogeneous microbial population consists of cells in different states, which implies a distribution of activities (e.g. respiration, product efficiency), as well as different responses to extracellular stimuli. The existence of a heterogeneous cell population may explain the lower productivities obtained for cultivations in large-scale reactors, where substrate and oxygen gradients are observed, in comparison to cultivations in well-mixed bench scale reactors (Enfors et al, 2001).

Population balance models (PBM) have been used in various applications (e.g. crystallization, granulation, flocculation, polymerization processes, etc.) to predict distributions of certain population properties including particle size, mass or volume, molecular weight,... Similarly, PBM allow for a mathematical description of distributed cell properties within microbial populations (Lencastre Fernandes et al, 2011).

Phenotypic heterogeneity arises as a result of the variability inherent to the metabolic mechanisms of single cells. Cell size is a key feature affecting cellular design, fitness and function (Jorgensen and Tyers, 2004). In fact, cell growth and division are tightly coupled, and this is reflected in the cell's capability of adjusting its growth rate to nutrients' availability (Enfors et al, 2001; Jorgensen and Tyers, 2004). Cell total protein content distributions (a measure of cell size) have been observed to provide a dynamic

picture of the interplay between the cells and surrounding extracellular environment (Alberghina et al, 1998).
The work hereby presented focuses on modeling the development of a microbial population growing in a continuous microbioreactor. The dynamics of single cell total protein content are described by a PBM, the consumption of substrate and production of ethanol are described by simple kinetics (unstructured model), and the transport of the species throughout the reactor is described by a computational fluid dynamics (CFD) model.

2. Case Study

The anaerobic fermentation of *Saccharomyces cerevisiae* in a continuous microreactor is used as case study. As above mentioned, cell growth and division are tightly coupled, and are modulated according to the substrate availability. The regulation of growth ensures that cells attain a critical size before initiating the division process (Rupeš, 2002; Porro et al, 1995). In the particular case of *S. cerevisiae*, two critical sizes corresponding to the regulation points START (committing to budding, or budding transition) and division have been identified. A schematic representation of the cells transition through the cell cycle and the associated cell growth (i.e. size increase) is presented in Fig. 1.
Using cell size as population descriptor allows, thus, for describing the distribution of cellular states. Moreover, a better description of the cellular state is obtained by determining the distribution of cells in cell cycle phases, i.e. by measuring DNA distributions (Alberghina et al, 1998). It was thus found desirable to use a PBM based on cell size as model variable, which is applied to different stages (i.e. subpopulations) corresponding to the non-budding and budding (cell cycle) phases.

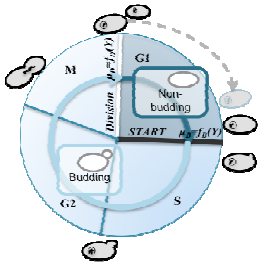

Figure 1 - Schematic representation of the cell cycle during exponential growth. The dark arrow (inside the circle) corresponds to the duration of the non-budding stage (G1phase). Due to their bigger size, G1 phase is shorter for mother cells than for daughter cells (dashed arrow). The *START* point represents the regulation point that defines the inititation of the DNA replication and budding process, i.e. transition to the S phase. The division point corresponds to the completion of the mitotic process and separation of a budding cell into two non-budding ones.

The microreactor consists of a sequence of spherical chambers connected in the center by a channel (Fig. 2). An inlet mass flow rate of $1x10^{-9}$ kg/s, containing glucose (20 g/l), was applied. The ethanol and biomass concentrations at the inlet are 0 g/l. The hydraulic retention time is 783 s.

3. Model Description

Aiming at both describing the distributions of single-cell sizes for the two cell size stages (non-budding and budding), a population balance model was coupled to an unstructured model describing the extracellular environment.

3.1. Population Balance Model

The PBM developed for this study is based on a multi-stage model reported in the literature (Hatzis and Porro, 2006). In this work, the segregated model consists of two

population balance equations (PBE), which describe cell growth, initiation of the budding process, cell division and birth, for both the non-budding and budding stages (Eq. 1 and 2), as well as the corresponding initial and boundary conditions.

$$\frac{\partial N^{NB}(m,t)}{\partial t}+\frac{\partial}{\partial m}\left[r_m(m,Y)N^{NB}(m,t)\right]=-\Gamma_B(m|Y)N^{NB}(m,t)+2\int_m^{\infty}\Gamma_D(m'|Y)P(m,m'|Y)N^B(m',t)dm' \quad \text{(Eq. 1)}$$

$$\frac{\partial N^{B}(m,t)}{\partial t}+\frac{\partial}{\partial m}\left[r_m(m,Y)N^{B}(m,t)\right]=-\Gamma_D(m|Y)N^{B}(m,t)+\Gamma_B(m|Y)N^{NB}(m,t) \quad \text{(Eq. 2)}$$

Here $N^{NB}dm$ and $N^{B}dm$ represent the number of cells in the cell size interval *[m, m+dm]* for the non-budding and budding stages, respectively, *m* is the cell total protein content or cell size (in arbitrary units), and *Y* designates the extracellular environment. The budding and division transitions, Γ_{NB} and Γ_B, are mathematically described as hazard functions (Eq. 3 and 4, for the budding transition) where the probability that a cell of mass *m* initiates the budding process, i.e. transitions to the budding stage, or divides into two new cells is described by a Gaussian probability density function with mean μ_B and μ_D, respectively. The standard deviation was assumed to be the same for the two transition functions and constant along the cultivation.

$$\Gamma_{NB}(m|Y)=r_m(m,Y)\frac{h_{NB}(m|Y)}{1-\int_{m_0}^{m}h_{NB}(m')dm'} \quad \text{(Eq. 3)}$$

$$h_{NB}(m|Y)=N(\mu_{NB}(Y),\sigma_{NB}) \quad \text{(Eq. 4)}$$

$$r_m(m,Y)=k_G m\cdot\lambda(Y) \quad \text{(Eq. 5)}$$

Generally, the link to the extracellular environment is accounted for by including substrate dependency in the growth function as well as transition functions (budding and division) for each of the stages: (i) the growth rate of single cells (r_m, *Eq. 5*) depends on the mass of each cell (first-order kinetics) and on the available concentration of glucose. The substrate factor, $\lambda(Y)$, can be regarded as a single cell specific growth rate and it is derived from the unstructured model describing the consumption of glucose and formation of ethanol; (ii) concerning the budding and division transitions, the critical transition cell sizes (μ_B and μ_D) are function of the locally available substrate concentration, and mathematically described by Monod type expressions. A mother and a daughter cell are generated upon division, where the ratio of the mother cell size to the daughter cell size is defined by the partitioning function, $P(m,m'|Y)$. This function consists of a symmetrical beta probability density function. The dependence of the partitioning function on the substrate was neglected for this case study.

3.2. Unstructured Model: Description of the extracellular environment

The unstructured model describes the fermentation of glucose (substrate) to ethanol. The local consumption rate is estimated based on an averaged yield of biomass on substrate and the concentration of biomass present at a given location in the microreactor and a given time point. The rate of formation of ethanol is estimated in a similar fashion. This concentration of biomass is calculated as the zeroth moment of the total cell size distribution. The updated concentrations of glucose and ethanol are used for recalculating the substrate factor $\lambda(Y)$, which serves as input to the PBM.

3.3. Solution methods

The fixed-pivot technique (Kumar and Ramkrishna, 1996) was used to discretize the PBE, using an evenly distributed grid with 20 pivots. The commercial software by ANSYS® CFX (v. 12.1) was used, and PBM was implemented using the expression language (CEL). The geometry was defined with ICEM CFD 12.1, and a hexahedral

mesh with 32159 elements and 36535 nodes was generated. Diffusion was not considered in the simulations.

4. Results and Discussion

The glucose supplied at the inlet is transported through the reactor, allowing for the biomass to grow. The biomass is suspended, and thus is also transported throughout the reactor until it exits the system. The spherical structures allow for a higher retention of the biomass within the reactor in comparison to a plug-flow configuration, preventing a complete wash out of the biomass. The steady-state profiles for glucose, total biomass and ethanol profiles are presented in Figure 2.

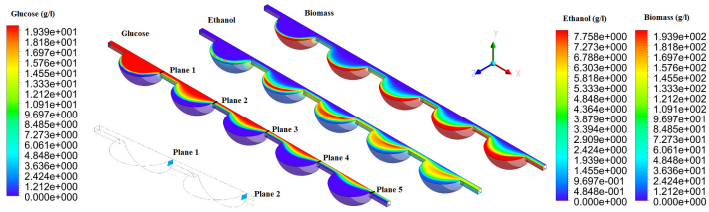

Figure 2 – Glucose, total biomass and ethanol profiles obtained at steady-state. As the reactor is symmetrical along the x- and y-axis, only a quarter of the reactor was simulated and illustrated. The total length of the reactor is 8 mm. The inlet of the reactor is located at the nearest to Plane 1.

The glucose concentration is highest at the inlet and decreases along the reactor. Therefore, the biomass growth rate is highest close to the inlet. The biomass formed in the first compartments is transported along the flow streamlines towards the outlet (Fig. 2). This results in a higher biomass concentration in the compartments closest to the outlet, although the growth rate here is virtually zero. Ethanol cannot be consumed by the cells due to the lack of oxygen. Its concentration is highest at the interface of glucose and the biomass is accumulated in each compartment.

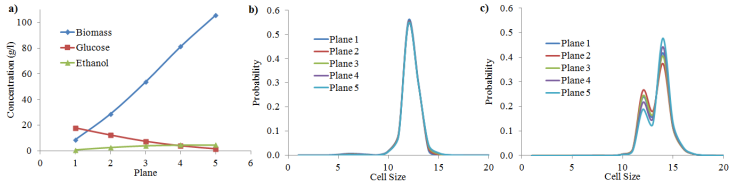

Figure 3 – Comparison of average concentrations and cell size probability density distributions for different location planes in the reactor: a) average concentrations of total biomass, glucose and ethanol b) distribution for the non-budding stage c) distribution for the budding stage. Arbitrary units are used for cell size.

Four different location planes were selected corresponding to the outlets of the spherical compartments (Fig.2) . The average glucose, ethanol and total biomass concentrations corresponding to each location plane are presented in the Figure 3. Additionally, the cell size distributions for the two stages for each of the locations are compared (Fig. 3). While the difference in the distributions is not significant for the non-budding stage, this difference is rather noticeable for the budding stage. If cells were imobilized in the

different compartments, it would be expected that the population in the compartment closest to the inlet (experiencing higher glucose concentration) would contain a higher fraction of bigger cells. Experimentally, the cell size distributions shifts towards smaller sizes when entering stationary phase (Werner-Washburne et al., 1993). This is however not observed in the simulation results, where cells are transported through the reactor. The coupling of a segregated biological model with the fluid dynamic model, increases the complexity of the problem and challenges the interpretation of the results. Different control tests, e.g. continuity check and mass balance closure, have been performed. Further investigation, both *in silico* and experimentally, is required in order to assess the validity of these results.

5. Conclusions

In this case study, a PBM describing the distribution of cell size for cell cycle stages, and its dependence on the available substrate concentration, was coupled to a fluid dynamic model that describes the transport of the supplied substrate and the biomass throughout the reactor. The interpretation of the results is challenging as the transport of the cells along the fluid streams has to be considered together with biological phenomena (growth, budding and division) taking place locally. The study is, nonetheless, a contribution to the development of modeling tools for successful prediction of the dynamic behavior of total protein content distributions of a yeast population under non-ideal mixing conditions, as found in large-scale fermentors as well. In the future a comparison of this more detailed model with simpler ones (e.g. lumped models) should be performed in order to identify the situations where the use of more complex models such as the proposed one is adequate and justifiable.

Acknowledgments

The Danish Council for Strategic Research is acknowledged for financial support in the frame of project number 09-065160.

References

S.-O. Enfors, M. Jahic, A. Rozkov et al. 2001. Physiological responses to mixing in large scale bioreactors. Journal of Biotechnology, 85, 175-185

R. Lencastre Fernandes, M. Nierychlo, L. Lundin et al. 2011. Experimental methods and modeling techniques for description of cell population heterogeneity. Biotechnology Advances, 29, 575-599

L. Alberghina, H.V. Westerhoff eds. 2005. Systems biology. Berlin/Heidelberg: Springer-Verlag. Vol. 13.

L. Alberghina, C. Smeraldi, B.M. Ranzi, D. Porro. 1998. Control by nutrients of growth and cell cycle progression in budding yeast, analyzed by double-tag flow cytometry. Journal of Bacteriology, 180, 3864-3872.

P. Jorgensen, I. Rupeš , J.R. Sharom et al. 2004. A dynamic transcriptional network communicates growth potential to ribosome synthesis and critical cell size. Genes & Development, 18, 2491 -2505.

I. Rupeš. 2002. Checking cell size in yeast. Trends in Genetics, 18, 479-485.

D. Porro, F. Srienc. 1995. Tracking of individual cell cohorts in asynchronous saccharomyces cerevisiae populations. Biotechnology Progress, 11, 342-347.

C. Hatzis, D. Porro. 2006. Morphologically-structured models of growing budding yeast populations. Journal of Biotechnology, 124, 420-438.

S. Kumar, D. Ramkrishna. 1996. On the solution of population balance equations by discretization—i. a fixed pivot technique. Chemical Engineering Science, 51,1311-1332.

M. Werner-Washburne, E. Braun, G.C. Johnston , R. A. Singer. 1993. Stationary phase in the yeast Saccharomyces cerevisiae. Microbiology and Molecular Biology Reviews, 57, 383-401

Appendix B

Population dynamics during batch cultivation

B.1 Budding Index Estimation: flow cytometry *vs.* microscopy

The estimations of the BI obtained by microscopic observation have been compared to the estimation based on DNA distribution used in this work. In a parallel experiment, triplicate cultivations in 500 ml shake flasks with a working volume of 100 ml were performed. The same strain was grown in the same medium as the cultivations reported in the manuscript. Two samples from the exponential growth phase on glucose (t = 13 h and 15 h) and two others from the growth phase on ethanol (t = 23 h and 24 h) were taken. The samples were taken in duplicates.

One duplicate was analyzed on the microscope (optical microscope, 40x objective). For each sample, two slides were prepared and 15 photo frames corresponding to different locations in the slide were recorded. The total number of cells and the number of budding cells was counted for each photo frame, and the aggregated fraction of budding cells in all the counted cells then yields the BI. In average, a total of 446 cells were counted per sample, and a minimum of 187 cells (total number of cells) was registered for sample t = 13 h for one of the replicates (SF3). The second sample duplicate was centrifuged, and cells were resuspended

in 70% ethanol. After a period of at least 48 h, the samples were split into two aliquots and the DNA staining procedure (Porro et al, 2003) was performed for the two aliquots. The DNA content was measured in the same flow cytometer mentioned in the manuscript. The BI was estimated following the procedure described in the manuscript. The estimated BI for each duplicate of the 4 samples is presented in Figure B.1.

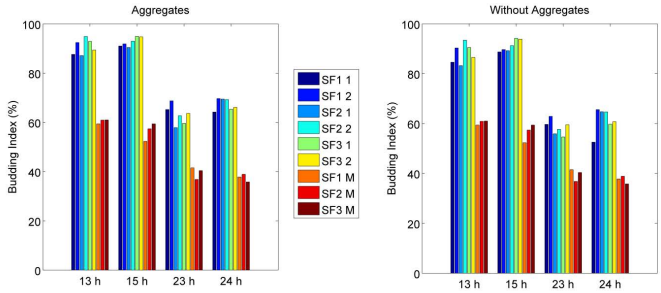

Figure B.1: Budding Index estimations based on analysis of the flow cytometric DNA histograms and microscopic cell counting: SFx identifies the triplicate shake flask cultivations, SFx 1 and 2 correspond to the two aliquots stained and analyzed in the flow cytometer, SFx M corresponds to the aggregated result for the microscopic cell count. The figure to the left (Aggregates) corresponds to a first analysis without excluding cell aggregates. In the figure to the right (Without Aggregates) events that presented a considerably high DNA content and FSC were disregarded for the flow cytometry based estimation of the BI.

Although the two methods (flow cytometry and microscopic) do not yield the same results, both methods show similar trends. Indeed, the decrease of the BI for the growth stage on ethanol in comparison to the initial growth on glucose is visible for both methods. In the case of the microscopic counting, when the bud is small cells may be incorrectly classified as non-budding, leading to an underestimation of the BI. On the contrary, in the case of the estimation based on flow cytometric data, in the presence of cell aggregates these will be classified as budding cells, despite the fact that the aggregates might consist of non-budding cells. In the case of microscopic counting, the samples were kept in an ultrasound bath in order to prevent aggregates. Aggregates observed were not taking into account in the counting procedure. Therefore, the estimation procedure based on flow cytometric DNA distributions may thus lead to an overestimation of the BI. We believe these are the reasons for the observed differences between the results obtained with the two methods.

When excluding aggregates (with a DNA > critical division band and high FSC), a bias in the flow cytometry based BI estimation was removed, as it can be observed in Figure B.2: the y-intercept is close to zero when the aggregates are removed before the BI estimation.

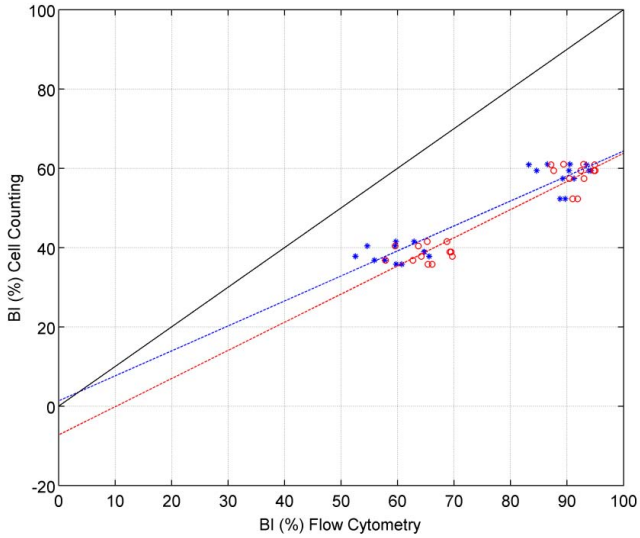

Figure B.2: Comparison of the BI estimation methods (microscopic cell counting vs. flow cytometry DNA histogram analysis): blue markers and line correspond to the analysis were aggregates are disregarded, while aggregates were considered in the results corresponding to red markers and line.

B.2 Cell total protein content: measurements and channel number as arbitrary unit

Experimentally the total protein content and the DNA content of a single cell can be determined by staining cells with fluorescein isothiocyanate and measure the fluorescence of each cell in a flow cytometer . However, the flow cytometer does not register an emission spectrum, or an absolute value for intensity, as in the

case of a fluorescent spectroscope. Instead, the fluorescence signal for an individual cell is recorded as a value on an arbitrary scale of channel numbers, corresponding to the bins for a histogram. Typically, flow cytometry histograms are presented in this arbitrary scale in a range of channel numbers from 0 to 1024 in the case of a linear scale, or from 100 to 104 in the case of a logarithmic scale. The first is generally used for forward and side scatter signals, whereas the latter is used for fluorescence signals. A calibration curve using fluorescence beads allows for translating the arbitrary scale of channel numbers into e.g. a scale of molecular equivalents of fluorochrome. In the case of both the total protein and DNA, the number of fluorochromes attaching to the cell is linearly proportional to the protein content or DNA content of an individual cell (calibration data not shown). For example, for a cell with two copies of the chromosome it can be assumed that the double amount of fluorochromes will attach to the DNA compared to a cell with only one copy of the chromosome. Similarly, as a high percentage of the cell content is made of protein, a higher number of fluorochromes will attach to the cell corresponding to a higher channel number in the arbitrary scale, than for a smaller cell. In this work, the channel number (ch. no.) arbitrary unit is used to describe the cell total protein content (i.e. cell size). Further details on the materials and methods used for collecting flow cytometric data within this study are provided in Part I of this work. A comprehensive description and discussion on flow cytometry can be found elsewhere [66].

B.3 Comparison of three replicate bioreactors

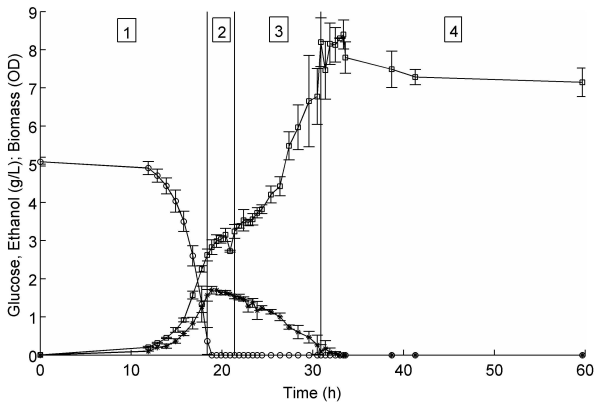

Figure B.3: Variation of glucose (open circles), ethanol (stars) and biomass (open squares) along the cultivation. The error bars correspond to the standard deviation of the three replicate cultivations. The numbers correspond to the different cultivation phases: 1) growth on glucose; 2) diauxic shift; 3) growth on ethanol; 4) stationary phase.

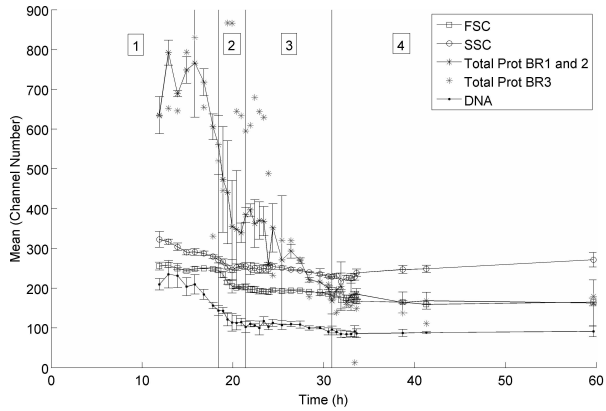

Figure B.4: Variation of the mean forward scatter (open squares), side scatter (open diamonds), total protein conten (stars) and DNA (dots). The error bars correspond to the standard deviation of the three replicate cultivations (marked as BR1, BR2 and BR3). In the case of the total protein content one of the reactors showed a deviant behavior during the diauxic shift and early growth on ethanol (star points, no line). This cultivation was disregarded when determining the mean values and standard deviations (stars and full line). The numbers correspond to the different cultivation phases: 1) growth on glucose; 2) diauxic shift; 3) growth on ethanol; 4) stationary phase..

B.4 Total Protein Content and DNA Content Histograms for all sample times and triplicate bioreactors

See next pages

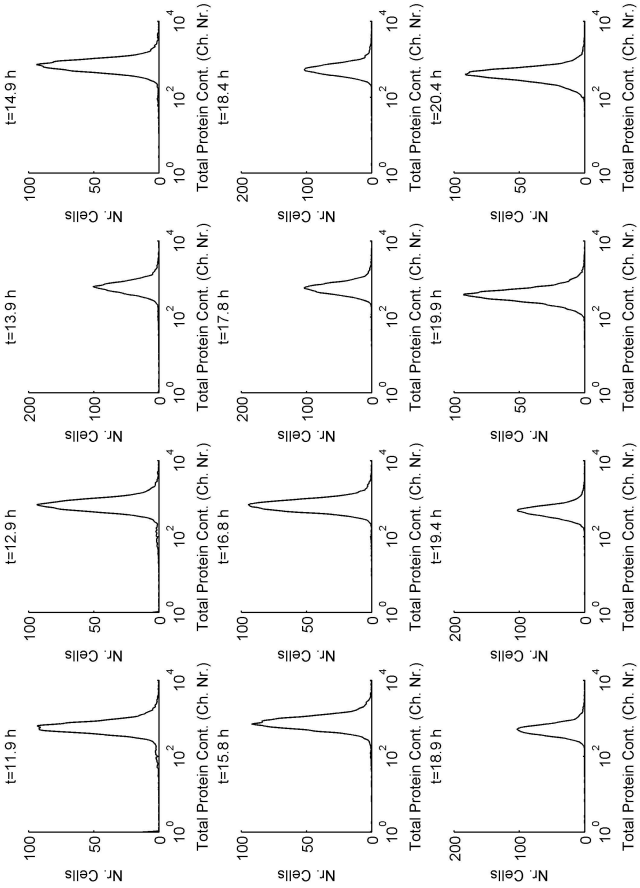

Figure B.5: Total protein content distribution for Bioreactor 1

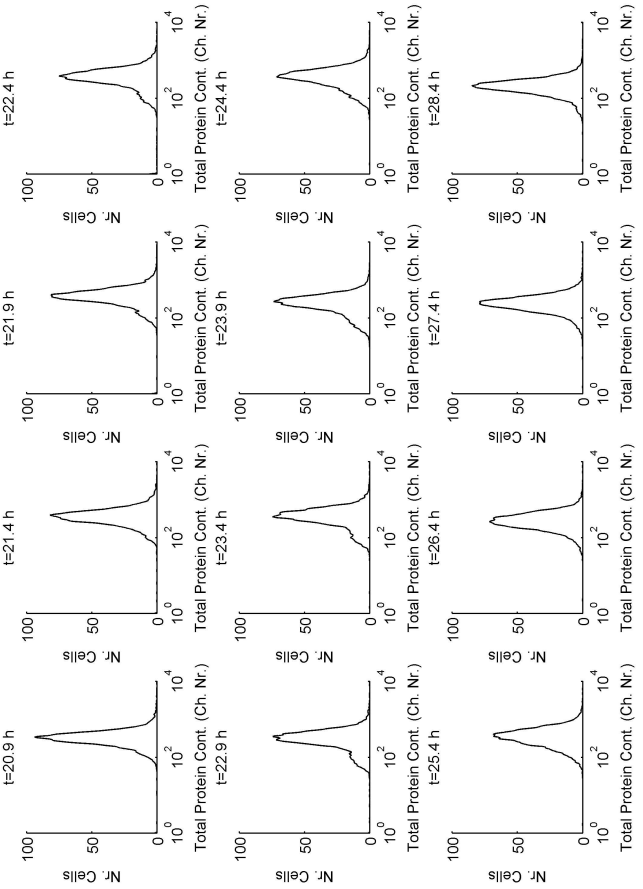

Figure B.6: Total protein content distribution for Bioreactor 2

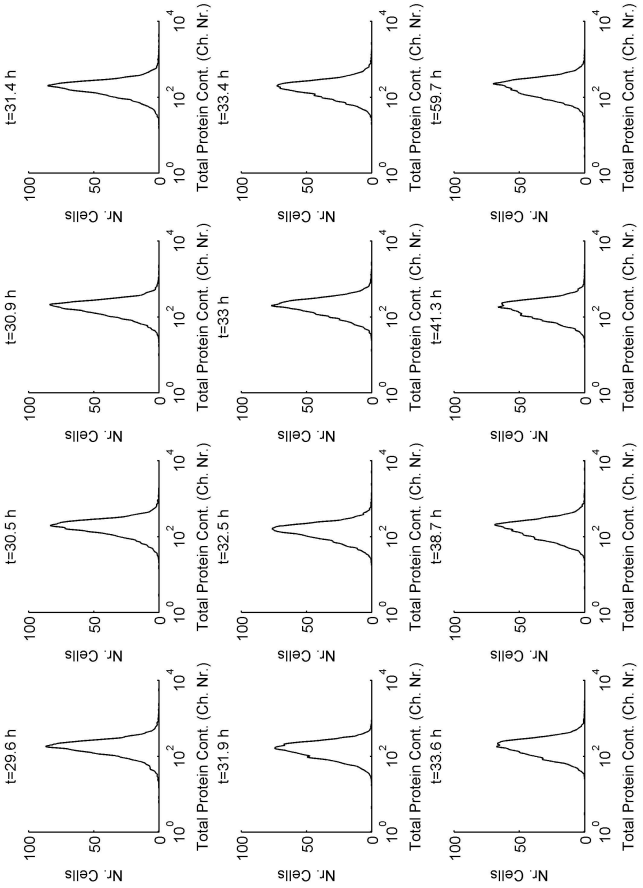

Figure B.7: Total protein content distribution for Bioreactor 3

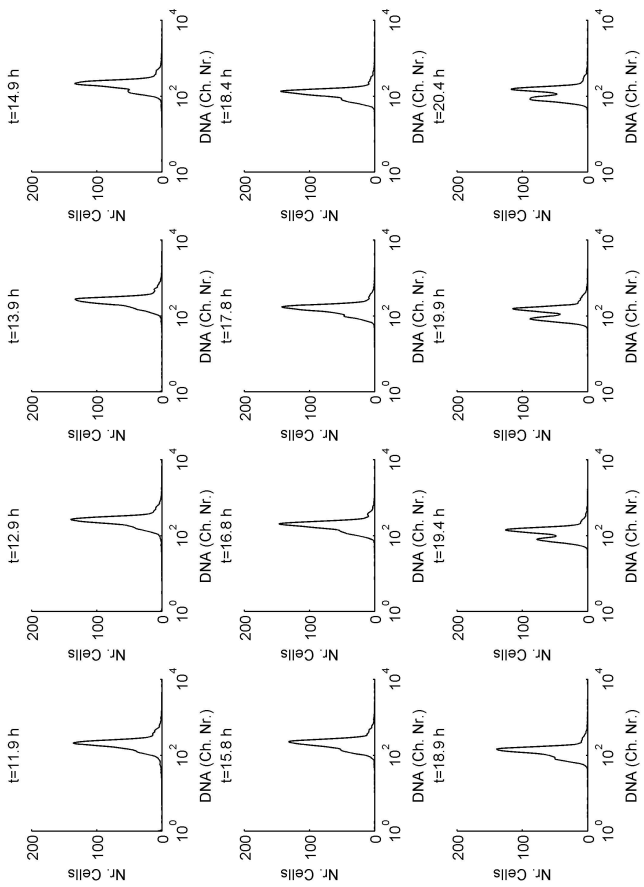

Figure B.8: DNA content distribution for Bioreactor 1

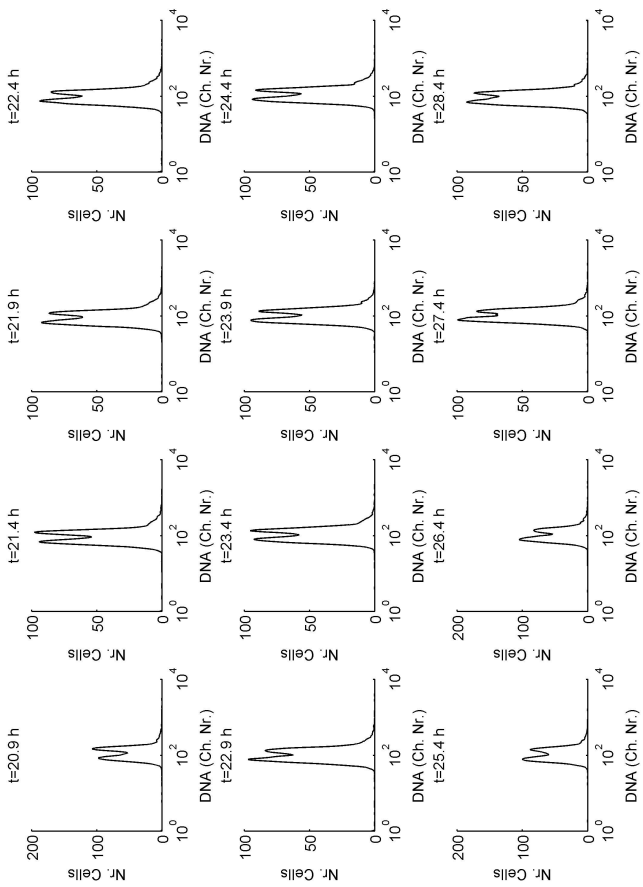

Figure B.9: DNA content distribution for Bioreactor 2

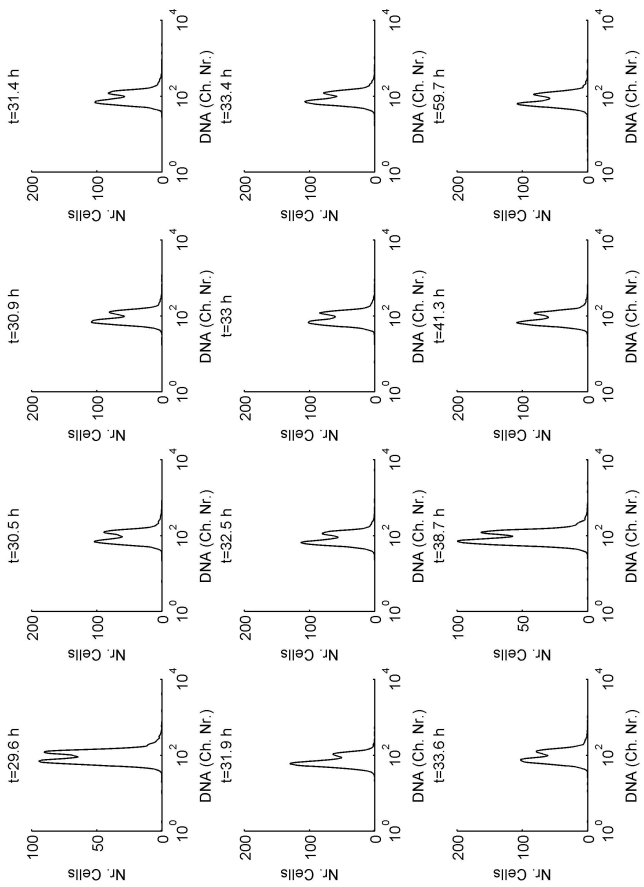

Figure B.10: DNA content distribution for Bioreactor 3

B.5 Comparison of beta distributions for various shape parameter combinations

The partitioning function is described by a beta distribution. The corresponding probability density function is defined by the following equation:

$$f\left(\frac{m}{m'};\alpha,\beta\right) = \frac{1}{B(\alpha,\beta)}\frac{m}{m'}^{\alpha-1}\left(1-\frac{m}{m'}\right)^{\beta-1} \tag{B.1}$$

where $B(\alpha,\beta)$ is a beta function, α and β are designated as shape parameters, and m/m' is the ratio of the size of the daughter cell originated as the bud to the size of original budded cell. In order to compare and interpret the model outputs obtained using for various combinations for the shape parameters, it is important to have in mind how the parameters modify the shape of the beta distribution (i.e. partitioning function).

The beta distributions for the parameter combinations used for the sensitivity analysis included in this work are illustrated in Figures B.11 and B.12.

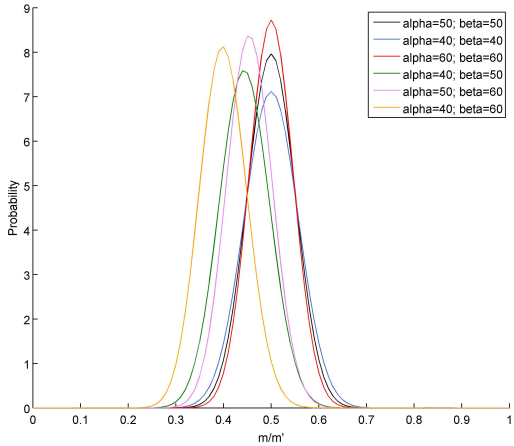

Figure B.11: Beta probability density functions for the shape parameter combinations used for Case I of the sensitivity analysis.

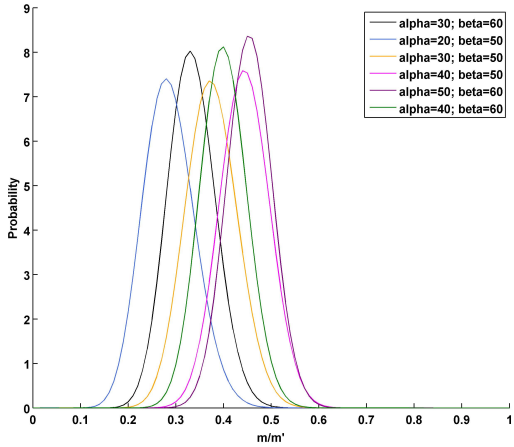

Figure B.12: Beta probability density functions for the shape parameter combinations used for Case II of the sensitivity analysis.

Bibliography

[1] Carlquist M., Lencastre Fernandes R., Helmark S., et al. Physiological heterogeneities in microbial populations and implications for physical stress tolerance. *Microbial Cell Factories*, 11(1):94, 2012.

[2] Bylund F., Collet E., Enfors S.-O., and Larsson G. Substrate gradient formation in the large-scale bioreactor lowers cell yield and increases by-product formation. *Bioprocess and Biosystems Engineering*, 18(3):171–180, 1998.

[3] Enfors S. O., Jahic M., Rozkov A., et al. Physiological responses to mixing in large scale bioreactors. *Journal of Biotechnology*, 85(2):175–185, 2001.

[4] George S., Larsson G., Olsson K., and Enfors S.-O. Comparison of the baker's yeast process performance in laboratory and production scale. *Bioprocess and Biosystems Engineering*, 18(2):135–142, 1998.

[5] Larsson G., Törnkvist M., Wernersson E. S. h., et al. Substrate gradients in bioreactors: origin and consequences. *Bioprocess and Biosystems Engineering*, 14(6):281–289, 1996.

[6] Kitano H. Cancer as a robust system: implications for anticancer therapy. *Nat Rev Cancer*, 4(3):227–235, 2004.

[7] Balaban N. Q., Merrin J., Chait R., Kowalik L., and Leibler S. Bacterial persistence as a phenotypic switch. *Science*, 305(5690):1622–1625, 2004.

[8] Papagianni M. Fungal morphology and metabolite production in submerged mycelial processes. *Biotechnology Advances*, 22(3):189–259, 2004.

[9] Dhar N. and McKinney J. D. Microbial phenotypic heterogeneity and antibiotic tolerance. *Current Opinion in Microbiology*, 10(1):30–38, 2007.

[10] Gernaey K. V., Lantz A. E., Tufvesson P., Woodley J. M., and Sin G. Application of mechanistic models to fermentation and biocatalysis for next-generation processes. *Trends in Biotechnology*, 28:346–354, 2010.

[11] Schmid A., Kortmann H., Dittrich P. S., and Blank L. M. Chemical and biological single cell analysis. *Current Opinion in Biotechnology*, 21(1):12–20, 2010.

[12] Müller S., Harms H., and Bley T. Origin and analysis of microbial population heterogeneity in bioprocesses. *Current Opinion in Biotechnology*, 21(1): 100–113, 2010.

[13] Fredrickson A. G., Megeee R. D., III, and Tsuchiya H. Mathematical models

in fermentation processes. *Advances in Applied Microbiology*, 13:419–465, 1970.

[14] Nopens I., Briesen H., and Ducoste J. Celebrating a milestone in population balance modeling. *Chemical Engineering Science*, 64(4):627–627, 2009.

[15] Karathia H., Vilaprinyo E., Sorribas A., and Alves R. *Saccharomyces cerevisiae* as a model organism: A comparative study. *PLoS ONE*, 6(2), 2011.

[16] Altmann K., Dürr M., and Westermann B. *Saccharomyces cerevisiae* as a model organism to study mitochondrial biology. In Leister D. and Herrmann J. M., editors, *Methods in Molecular Biology*, volume 372, pages 81–90. Humana Press, 2007.

[17] Drubin D. The yeast *Saccharomyces cerevisiae* as a model organism for the cytoskeleton and cellbiology. *Cell Motility and the Cytoskeleton*, 14(1): 42–49, 1989.

[18] Goffeau A., Barrell B. G., Bussey H., et al. Life with 6000 genes. *Science (New York, N.Y.)*, 274(5287):563–567, 1996.

[19] Venter J. C., Adams M. D., Myers E. W., et al. The sequence of the human genome. *Science*, 291(5507):1304–1351, 2001.

[20] Cherry J. M., Adler C., Ball C., et al. Sgd: Saccharomyces genome database. *Nucleic Acids Research*, 26(1):73–79, 1998.

[21] Nissen T. L., Kielland-Brandt M. C., Nielsen J., and Villadsen J. Optimization of ethanol production in *Saccharomyces cerevisiae* by metabolic engineering of the ammonium assimilation. *Metabolic Engineering*, 2(1):69–77, 2000.

[22] Ferrer-Miralles N., Domingo-Espín J., Corchero J., Vázquez E., and Villaverde A. Microbial factories for recombinant pharmaceuticals. *Microbial Cell Factories*, 8(1), 2009.

[23] Gerngross T. U. Advances in the production of human therapeutic proteins in yeasts and filamentous fungi. *Nature Biotechnology*, 22(11):1409–1414, 2004.

[24] Walsh G. Biopharmaceutical benchmarks 2010. *Nature Biotechnology*, 28 (9):917–924, 2010.

[25] Walsh G. New biopharmaceuticals: A review of new biologic drug approvals over the years, featuring highlights from 2010 and 2011. Technical report, Process Development Forum, 2012.

[26] Martínez J. L., Liu L., Petranovic D., and Nielsen J. Pharmaceutical protein production by yeast: towards production of human blood proteins by microbial fermentation. *Current Opinion in Biotechnology*, Available online, 2012.

[27] Roldão A., Kim I.-K., and Nielsen J. Bridging omics technologies with synthetic biology in yeast industrial biotechnology. In Wittmann C. and Lee S. Y., editors, *Systems Metabolic Engineering*, pages 271–327. Springer Netherlands, 2012.

[28] Le Borgne S. Genetic engineering of industrial strains of *Saccharomyces*

cerevisiae. *Methods in molecular biology (Clifton, N.J.)*, 824:451–465, 2012.

[29] Singhania R. R., Patel A. K., Soccol C. R., and Pandey A. Recent advances in solid-state fermentation. *Biochemical Engineering Journal*, 44(1):13–18, 2009.

[30] Martins S., Mussatto S. I., Martínez-Avila G., et al. Bioactive phenolic compounds: Production and extraction by solid-state fermentation. a review. *Biotechnology Advances*, 29(3):365–373, 2011.

[31] Chen L. Z., Nguang S. K., and Chen X. D. *Modelling and Optimization of Biotechnological Processes: Artificial Intelligence Approaches*. Springer Berlin / Heidelberg, 2006.

[32] Hensing M. C. M., Rouwenhorst R. J., Heijnen J. J., Dijken J. P., and Pronk J. T. Physiological and technological aspects of large-scale heterologous-protein production with yeasts. *Antonie van Leeuwenhoek*, 67(3):261–279, 1995.

[33] Sonnleitner B. Instrumentation of biotechnological processes. In Sonnleitner B., editor, *Advances in Biochemical Engineering/Biotechnology*, volume 66, pages 1–64. Springer Berlin / Heidelberg, 2000.

[34] Nielsen J. and Villadsen J. Modelling of microbial kinetics. *Chemical Engineering Science*, 47(17-18):4225–4270, 1992.

[35] Agger T., Spohr A. B., Carlsen M., and Nielsen J. Growth and product formation of *Aspergillus oryzae* during submerged cultivations: verification of a morphologically structured model using fluorescent probes. *Biotechnology and Bioengineering*, 57(3):321–329, 1998.

[36] Jones K. D. and Kompala D. S. Cybernetic model of the growth dynamics of *Saccharomyces cerevisiae* in batch and continuous cultures. *Journal of Biotechnology*, 71(1-3):105–131, 1999.

[37] Young J., Henne K., Morgan J., Konopka A., and Ramkrishna D. Cybernetic modeling of metabolism: towards a framework for rational design of recombinant organisms. *Chemical Engineering Science*, 59(22-23):5041–5049, 2004.

[38] Becker S. A., Feist A. M., Mo M. L., et al. Quantitative prediction of cellular metabolism with constraint-based models: the cobra toolbox. *Nature Protocols*, 2(3):727–738, 2007.

[39] Gombert A. K. and Nielsen J. Mathematical modelling of metabolism. *Current Opinion in Biotechnology*, 11(2):180–186, 2000.

[40] Jong H.de . Modeling and simulation of genetic regulatory systems: A literature review. *Journal of Computational Biology*, 9(1):67–103, 2002.

[41] Stelling J. Mathematical models in microbial systems biology. *Current Opinion in Microbiology*, 7(5):513–518, 2004.

[42] Teusink B. and Smid E. J. Modelling strategies for the industrial exploitation of lactic acid bacteria. *Nature Reviews Microbiology*, 4(1):46–56, 2006.

[43] Aertsen A. and Michiels C. W. Stress and how bacteria cope with death and survival. *Critical Reviews in Microbiology*, 30(4):–, 2004.

[44] Booth I. R. Stress and the single cell: Intrapopulation diversity is a mechanism to ensure survival upon exposure to stress. *International Journal of Food Microbiology*, 78(1-2):19–30, 2002.

[45] Lengeler J., Drews G., and Schlegel H. *Biology of Prokaryotes*. Thieme, New York, 1999.

[46] Norman A., Hansen L. H., and Sørensen S. J. A flow cytometry-optimized assay using an *SOS*-green fluorescent protein (*SOS-GFP*) whole-cell biosensor for the detection of genotoxins in complex environments. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, 603(2):164–172, 2006.

[47] Ron E. Bacterial stress response. In *The Prokaryotes*, pages 1012–1027. 2006.

[48] Marles-Wright J. and Lewis R. J. Stress responses of bacteria. *Current Opinion in Structural Biology*, 17(6):755–760, 2007.

[49] De Angelis M., Bini L., Pallini V., Cocconcelli P. S., and Gobbetti M. The acid-stress response in lactobacillus sanfranciscensis cb1. *Microbiology*, 147(7): 1863–1873, 2001.

[50] Sørensen S. J., Burmølle M., and Hansen L. H. Making bio-sense of toxicity: new developments in whole-cell biosensors. *Current Opinion in Biotechnology*, 17(1):11–16, 2006.

[51] Serrazanetti D. I., Guerzoni M. E., Corsetti A., and Vogel R. Metabolic impact and potential exploitation of the stress reactions in *Lactobacilli*. *Food Microbiology*, 26(7):700–711, 2009.

[52] Fritzsch F. S., Dusny C., Frick O., and Schmid A. Single-cell analysis in biotechnology, systems biology, and biocatalysis. *Annual Review of Chemical and Biomolecular Engineering*, 3(1):129–155, 2012.

[53] Brehm-Stecher B. F. and Johnson E. A. Single-cell microbiology: Tools, technologies, and applications. *Microbiology and Molecular Biology Reviews*, 68 (3):538–559, 2004.

[54] Mandy F. F., Bergeron M., and Minkus T. Principles of flow cytometry. *Transfusion Science*, 16(4):303–314, 1995.

[55] Rieseberg M., Kasper C., Reardon K. F., and Scheper T. Flow cytometry in biotechnology. *Applied Microbiology and Biotechnology*, 56(3):350–360, 2001.

[56] Davey H. and Kell D. Flow cytometry and cell sorting of heterogeneous microbial populations: the importance of single-cell analyses. *Microbiology Reviews*, 60(4):641–696, 1996.

[57] Díaz M., Herrero M., García L. A., and Quirós C. Application of flow cytometry to industrial microbial bioprocesses. *Biochemical Engineering Journal*, 48 (3):385–407, 2010.

[58] Mateus C. and Avery S. V. Destabilized green fluorescent protein for monitoring dynamic changes in yeast gene expression with flow cytometry.

Yeast, 16(14):1313–1323, 2000.

[59] Kacmar J., Gilbert A., Cockrell J., and Srienc F. The cytostat: A new way to study cell physiology in a precisely defined environment. *Journal of Biotechnology*, 126(2):163–172, 2006.

[60] Sørensen S. J., Sørensen A. H., Hansen L. H., Oregaard G., and Veal D. Direct detection and quantification of horizontal gene transfer by using flow cytometry and gfp as a reporter gene. *Current Microbiology*, 47(2):129–133, 2003.

[61] Attfield P., Kletsas S., Veal D., Van Rooijen R., and Bell P. Use of flow cytometry to monitor cell damage and predict fermentation activity of dried yeasts. *Journal of Applied Microbiology*, 89(2):207–214, 2000.

[62] Davey H. M. and Winson M. K. Using flow cytometry to quantify microbial heterogeneity. *Current issues in molecular biology*, 5(1):9–15, 2003.

[63] Bergquist P., Hardiman E., Ferrari B., and Winsley T. Applications of flow cytometry in environmental microbiology and biotechnology. *Extremophiles*, 13(3):389–401, 2009.

[64] Czechowska K., Johnson D. R., and Meer J. R.van der . Use of flow cytometric methods for single-cell analysis in environmental microbiology. *Current Opinion in Microbiology*, 11(3):205–212, 2008.

[65] Link A. J., Jeong K. J., and Georgiou G. Beyond toothpicks: new methods for isolating mutant bacteria. *Nature Reviews Microbiology*, 5(9):680–688, 2007.

[66] Shapiro H. M. *Practical flow cytometry*. John Wiley and Sons, 2004.

[67] Ishii S., Tago K., and Senoo K. Single-cell analysis and isolation for microbiology and biotechnology: methods and applications. *Applied Microbiology and Biotechnology*, 86(5):1281–1292, 2010.

[68] Caron G.Nebe-von , Stephens P. J., Hewitt C. J., Powell J. R., and Badley R. A. Analysis of bacterial function by multi-colour fluorescence flow cytometry and single cell sorting. *Journal of Microbiological Methods*, 42(1):97–114, 2000.

[69] Allegra S., Berger F., Berthelot P., et al. Use of flow cytometry to monitor legionella viability. *Applied and Environmental Microbiology*, 74(24):7813–7816, 2008.

[70] Ben-Amor K., Heilig H., Smidt H., et al. Genetic diversity of viable, injured, and dead fecal bacteria assessed by fluorescence-activated cell sorting and *16S rRNA* gene analysis. *Applied Environmental Microbiology*, 71(8):4679–4689, 2005.

[71] Berney M., Hammes F., Bosshard F., Weilenmann H.-U., and Egli T. Assessment and interpretation of bacterial viability by using the live/dead baclight kit in combination with flow cytometry. *Applied Environmental Microbiology*, 73(10):3283–3290, 2007.

[72] Nielsen T. H., Sjøholm O. R., and Sørensen J. Multiple physiological states of

a *Pseudomonas fluorescens* dr54 biocontrol inoculant monitored by a new flow cytometry protocol. *FEMS Microbiology Ecology*, 67(3):479–490, 2009.

[73] Novo D., Perlmutter N. G., Hunt R. H., and Shapiro H. M. Accurate flow cytometric membrane potential measurement in bacteria using diethyloxacarbocyanine and a ratiometric technique. *Cytometry*, 35(1):55–63, 1999.

[74] Müller S., Ullrich S., Lsche A., Loffhagen N., and Babel W. Flow cytometric techniques to characterise physiological states of *Acinetobacter calcoaceticus*. *Journal of Microbiological Methods*, 40(1):67–77, 2000.

[75] Falcioni T., Papa S., and Gasol J. M. Evaluating the flow-cytometric nucleic acid double-staining protocol in realistic situations of planktonic bacterial death. *Applied Environmental Microbiology*, 74(6):1767–1779, 2008.

[76] Gasol J. M. and Arstegui J. Cytometric evidence reconciling the toxicity and usefulness of ctc as a marker of bacterial activity. *Aquatic Microbial Ecology*, 46(1):71–83, 2007.

[77] Caro A., Gros O., Got P., De Wit R., and Troussellier M. Characterization of the population of the sulfur-oxidizing symbiont of *Codakia orbicularis* (bivalvia, *Lucinidae*) by single-cell analyses. *Applied Environmental Microbiology*, 73 (7):2101–2109, 2007.

[78] Hoefel D., Grooby W. L., Monis P. T., Andrews S., and Saint C. P. Enumeration of water-borne bacteria using viability assays and flow cytometry: a comparison to culture-based techniques. *Journal of Microbiological Methods*, 55 (3):585–597, 2003.

[79] Bunthof C. J., Braak S.van den , Breeuwer P., Rombouts F. M., and Abee T. Rapid fluorescence assessment of the viability of stressed lactococcus lactis. *Applied Environmental Microbiology*, 65(8):3681–3689, 1999.

[80] Cronin U. and Wilkinson M. Monitoring growth phase-related changes in phosphatidylcholine-specific phospholipase c production, adhesion properties and physiology of bacillus cereus vegetative cells. *Journal of Industrial Microbiology and Biotechnology*, 35(12):1695–1703, 2008.

[81] Rault A., Bouix M., and Beal C. Fermentation ph influences the physiological-state dynamics of lactobacillus bulgaricus cfl1 during ph-controlled culture. *Applied Environmental Microbiology*, 75(13):4374–4381, 2009.

[82] Herzenberg L. A., Tung J., Moore W. A., Herzenberg L. A., and Parks D. R. Interpreting flow cytometry data: a guide for the perplexed. *Nature immunology*, 7(7):681–685, 2006.

[83] Bahl M. I., Sørensen S. J., and Hestbjerg Hansen L. Quantification of plasmid loss in *Escherichia coli* cells by use of flow cytometry. *FEMS Microbiology Letters*, 232(1):45–49, 2004.

[84] Hammes F. and Egli T. Cytometric methods for measuring bacteria in water: advantages, pitfalls and applications. *Analytical and bioanalytical chemistry*, 397(3):1083–1095, 2010.

[85] Spidlen J., Gentleman R. C., Haaland P. D., et al. Data standards for flow

cytometry. *OMICS: A Journal of Integrative Biology*, 10(2):209–214, 2006.

[86] Wang Y., Hammes F., De Roy K., Verstraete W., and Boon N. Past, present and future applications of flow cytometry in aquatic microbiology. *Trends in Biotechnology*, 28(8):416–424, 2010.

[87] De Rosa S. C. Vaccine applications of flow cytometry. *Methods*, 57(3):383–391, 2012.

[88] Silva T. L. d., Roseiro J. C., and Reis A. Applications and perspectives of multi-parameter flow cytometry to microbial biofuels production processes. *Trends in Biotechnology*, 30(4):225–232, 2012.

[89] Spidlen J., Leif R. C., Moore W., Roederer M., and Brinkman R. R. Gating-ml: Xml-based gating descriptions in flow cytometry. *Cytometry Part A*, 73A (12):1151–1157, 2008.

[90] Lo K., Brinkman R. R., and Gottardo R. Automated gating of flow cytometry data via robust model-based clustering. *Cytometry Part A*, 73A(4):321–332, 2008.

[91] Caron G.Nebe-von . Standardization in microbial cytometry. *Cytometry Part A*, 75A(2):86–89, 2009.

[92] Lee J. A., Spidlen J., Boyce K., et al. Miflowcyt: The minimum information about a flow cytometry experiment. *Cytometry Part A*, 73A(10):926–930, 2008.

[93] Pyne S., Hu X., Wang K., et al. Automated high-dimensional flow cytometric data analysis. *Proceedings of the National Academy of Sciences*, 106(21): 8519–8524, 2009.

[94] Bashashati A. and Brinkman R. R. A survey of flow cytometry data analysis methods. *Advances in Bioinformatics*, 2009:1–19, 2009.

[95] Davey H. M. Prospects for the automation of analysis and interpretation of flow cytometric data. *Cytometry Part A*, 77A(1):3–5, 2010.

[96] Hahne F., LeMeur N., Brinkman R. R., et al. flowcore: a bioconductor package for high throughput flow cytometry. *BMC Bioinformatics*, 10(1), 2009.

[97] Lo K., Hahne F., Brinkman R. R., and Gottardo R. flowclust: a bioconductor package for automated gating of flow cytometry data. *BMC Bioinformatics*, 10(1), 2009.

[98] Broger T., Odermatt R. P., Huber P., and Sonnleitner B. Real-time on-line flow cytometry for bioprocess monitoring. *Journal of Biotechnology*, 154 (4):240–247, 2011.

[99] Fišer K., Sieger T., Schumich A., et al. Detection and monitoring of normal and leukemic cell populations with hierarchical clustering of flow cytometry data. *Cytometry Part A*, 81A(1):25–34, 2012.

[100] Fan Z.-C., Yan J., Liu G.-D., et al. Real-time monitoring of rare circulating hepatocellular carcinoma cells in an orthotopic model by in vivo flow cytometry assesses resection on metastasis. *Cancer Research*, 72(10):2683–2691, 2012.

[101] Boedigheimer M. J. and Ferbas J. Mixture modeling approach to flow cytometry data. *Cytometry Part A*, 73A(5):421–429, 2008.

[102] Ramkrishna D. On modeling of bioreactors for control. *Journal of Process Control*, 13(7):581–589, 2003.

[103] Daoutidis P. and Henson M. Dynamics and control of cell populations in continuous bioreactors. *AiChE Symposium Series*, 326:393–402, 2002.

[104] Fredrickson A. G. Population balance equations for cell and microbial cultures revisited. *AIChE Journal*, 49(4):1050–1059, 2003.

[105] Ramkrishna D. *Population Balances: Theory and Applications to Particulate Systems in Engineering*. Academic Press, 2000.

[106] Bailey J. E. Mathematical modeling and analysis in biochemical engineering: past accomplishments and future opportunities. *Biotechnology Progress*, 14(1):8–20, 1998.

[107] Gernaey K. V., Loosdrecht M. C.van , Henze M., Lind M., and Jørgensen S. B. Activated sludge wastewater treatment plant modelling and simulation: state of the art. *Environmental Modelling & Software*, 19(9):763–783, 2004.

[108] Henze M., Gujer W., Mino T., and Loosedrecht M.van . *Activated Sludge Models ASM1, ASM2, ASM2d and ASM3*. IWA Publishing, 2006.

[109] Cazzador L. and Mariani L. A simulation program based on a structured population model for biotechnological yeast processes. *Applied Microbiology and Biotechnology*, 29(2):198–202, 1988.

[110] Hatzis C. and Porro D. Morphologically-structured models of growing budding yeast populations. *Journal of Biotechnology*, 124(2):420–438, 2006.

[111] Mantzaris N. V., Srienc F., and Daoutidis P. Nonlinear productivity control using a multi-staged cell population balance model. *Chemical Engineering Science*, 57(1):1–14, 2002.

[112] Bley T. From single cells to microbial population dynamics: modelling in biotechnology based on measurements of individual cells. *Advances in biochemical engineering/biotechnology*, 124:211–227, 2011.

[113] Mantzaris N. V., Daoutidis P., and Srienc F. Numerical solution of multivariable cell population balance models: I. finite difference methods. *Computers & Chemical Engineering*, 25(11-12):1411–1440, 2001.

[114] Henson M. A. Dynamic modeling of microbial cell populations. *Current Opinion in Biotechnology*, 14(5):460–467, 2003.

[115] Henson M. A. Modeling the synchronization of yeast respiratory oscillations. *Journal of Theoretical Biology*, 231(3):443–458, 2004.

[116] Henson M. A. Cell ensemble modeling of metabolic oscillations in continuous yeast cultures. *Computers & Chemical Engineering*, 29(3):645–661, 2005.

[117] Henson M. A., Müller D., and Reuss M. Cell population modelling of yeast glycolytic oscillations. *Biochemical Journal*, 368(Pt 2):433–446, 2002.

[118] Stamatakis M. Cell population balance, ensemble and continuum modeling frameworks: Conditional equivalence and hybrid approaches. *Chemical Engineering Science*, 65(2):1008–1015, 2010.

[119] Hjortsø M. A., M. and Nielsen J. Population balance models of autonomous microbial oscillations. *Journal of Biotechnology*, 42(3):255–269, 1995.

[120] Zamamiri A. M., Zhang Y., Henson M. A., and Hjortsø M. A. Dynamics analysis of an age distribution model of oscillating yeast cultures. *Chemical Engineering Science*, 57(12):2169–2181, 2002.

[121] Bellgardt K. H. Analysis of synchronous growth of baker's yeast. part i: Development of a theoretical model for sustained oscillations. *Journal of Biotechnology*, 35(1):19–33, 1994.

[122] Bellgardt K. H. Analysis of synchronous growth of baker's yeast. part ii: Comparison of model prediction and experimental data. *Journal of Biotechnology*, 35(1):35–49, 1994.

[123] Kurtz M. J., Zhu G.-Y., Zamamiri A., Henson M. A., and Hjortsø M. A. Control of oscillating microbial cultures described by population balance models. *Industrial & Engineering Chemistry Research*, 37(10):4059–4070, 1998.

[124] Duarte M., Medeiros J., Araújo O., and Coelho M. An age-structured population balance model for microbial dynamics. *Brazilian Journal of Chemical Engineering*, 20(1):–, 2003.

[125] Sherer E., Tocce E., Hannemann R., Rundell A., and Ramkrishna D. Identification of age-structured models: Cell cycle phase transitions. *Biotechnology and Bioengineering*, 99(4):960–974, 2008.

[126] Huang H., Ridgway D., Gu T., and Moo-Young M. A segregated model for heterologous amylase production by *Bacillus subtilis*. *Enzyme and Microbial Technology*, 32(3-4):407–413, 2003.

[127] Takamatsu T., Shioya S., Chikatani H., and Dairaku K. Comparison of simple population models in a baker's yeast fed-batch culture. *Chemical Engineering Science*, 40(3):499–507, 1985.

[128] Godin F. B., Cooper D. G., and Rey A. D. Numerical methods for a population-balance model of a periodic fermentation process. *AIChE Journal*, 45(6): 1359–1364, 1999.

[129] Mantzaris N. V., Liou J.-J., Daoutidis P., and Srienc F. Numerical solution of a mass structured cell population balance model in an environment of changing substrate concentration. *Journal of Biotechnology*, 71(1-3):157–174, 1999.

[130] Mhaskar P., Hjortsø M. A., and Henson M. A. Cell population modeling and parameter estimation for continuous cultures of *Saccharomyces cerevisiae*. *Biotechnology Progress*, 18(5):1010–1026, 2002.

[131] Subramanian G., Ramkrishna D., Fredrickson A., and Tsuchiya H. On the mass distribution model for microbial cell populations. *Bulletin of Mathematical Biology*, 32(4):521–537, 1970.

[132] Zhu G.-Y., Zamamiri A., Henson M. A., and Hjortsø M. A. Model predictive control of continuous yeast bioreactors using cell population balance models. *Chemical Engineering Science*, 55(24):6155–6167, 2000.

[133] Beuse M., Bartling R., Kopmann A., Diekmann H., and Thoma M. Effect of the dilution rate on the mode of oscillation in continuous cultures of *Saccharomyces cerevisiae*. *Journal of Biotechnology*, 61(1):15–31, 1998.

[134] Duboc P. and Stockar U.von . Modeling of oscillating cultivations of *Saccharomyces cerevisiae*: Identification of population structure and expansion kinetics based on on-line measurements. *Chemical Engineering Science*, 55(1):149–160, 2000.

[135] Cipollina C., Vai M., Porro D., and Hatzis C. Towards understanding of the complex structure of growing yeast populations. *Journal of Biotechnology*, 128(2):393–402, 2007.

[136] Hjortsø M. A. and Bailey J. E. Transient responses of budding yeast populations. *Mathematical Biosciences*, 63(1):121–148, 1983.

[137] Cazzador L., Mariani L., Martegani E., and Alberghina L. Structured segregated models and analysis of self-oscillating yeast continuous cultures. *Bioprocess and Biosystems Engineering*, 5(4):175–180, 1990.

[138] Hatzis C., Srienc F., and Fredrickson A. G. Multistaged corpuscular models of microbial growth: *Monte Carlo* simulations. *Biosystems*, 36(1):19–35, 1995.

[139] Fiolitakis E. Ein altersstrukturiertes populationsmodell zur beschreibung instationärer mikrobieller prozesse, teil ii: Modell-verifikation am beispiel der glucose-fermentation mit *Zymomonas mobilis*. *Chemie Ingenieur Technik*, 59(10):812–813, 1987.

[140] Fiolitakis E. Ein altersstrukturiertes populationsmodell zur beschreibung instationärer mikrobieller prozesse, teil i: Theorie. *Chemie Ingenieur Technik*, 59(10):810–811, 1987.

[141] Mariani L., Martegani E., and Alberghina L. Yeast population models for monitoring and control of biotechnical processes. *IEE Proceedings*, 133(5): 210–216, 1986.

[142] Liou J.-J., Srienc F., and Fredrickson A. G. Solutions of population balance models based on a successive generations approach. *Chemical Engineering Science*, 52(9):1529–1540, 1997.

[143] Fredrickson A., Ramkrishna D., and Tsuchiya H. Statistics and dynamics of procaryotic cell populations. *Mathematical Biosciences*, 1(3):327–374, 1967.

[144] Ramkrishna D. Toward a self-similar theory of microbial populations. *Biotechnology and Bioengineering*, 43(2):138–148, 1994.

[145] Mantzaris N. V. and Daoutidis P. Cell population balance modeling and control in continuous bioreactors. *Journal of Process Control*, 14(7):775–784, 2004.

[146] Fredrickson A. G. and Mantzaris N. V. A new set of population balance equa-

tions for microbial and cell cultures. *Chemical Engineering Science*, 57(12): 2265–2278, 2002.

[147] Hjortsø M. Solution and properties of age population balance models which assume discrete division ages. *Journal of Biotechnology*, 42(3):271–280, 1995.

[148] Alberghina L., Smeraldi C., Ranzi B. M., and Porro D. Control by nutrients of growth and cell cycle progression in budding yeast, analyzed by double-tag flow cytometry. *Journal of Bacteriology*, 180(15):3864–3872, 1998.

[149] Porro D., Vai M., Vanoni M., Alberghina L., and Hatzis C. Analysis and modeling of growing budding yeast populations at the single cell level. *Cytometry Part A*, 75A(2):114–120, 2009.

[150] Wheals A. E. Size control models of *Saccharomyces cerevisiae* cell proliferation. *Molecular and Cellular Biology*, 2(4):361–368, 1982.

[151] Mantzaris N. V., Daoutidis P., and Srienc F. Numerical solution of multivariable cell population balance models. ii. spectral methods. *Computers & Chemical Engineering*, 25(11-12):1441–1462, 2001.

[152] Mantzaris N. V., Daoutidis P., and Srienc F. Numerical solution of multivariable cell population balance models. iii. finite element methods. *Computers & Chemical Engineering*, 25(11-12):1463–1481, 2001.

[153] Pollard , Ison , Shamlou , and Lilly . Reactor heterogeneity with *Saccharopolyspora erythraea* airlift fermentations. *Biotechnology and Bioengineering*, 58(5):453–463, 1998.

[154] Lapin A., Muller D., and Reuss M. Dynamic behavior of microbial populations in stirred bioreactors simulated with *Euler-Lagrange* methods: Traveling along the lifelines of single cells. *Industrial & Engineering Chemistry Research*, 43(16):4647–4656, 2004.

[155] Bezzo F., Macchietto S., and Pantelides C. C. Computational issues in hybrid multizonal/computational fluid dynamics models. *AIChE Journal*, 51(4): 1169–1177, 2005.

[156] Fang Z. D. Applying computational fluid dynamics technology in bioprocesses - part 1. *BioPharm International*, 23(4):38–45, 2010.

[157] Fang Z. D. Applying computational fluid dynamics technology in bioprocesses - part 2. *BioPharm International*, 23(5):42–45, 2010.

[158] Generalis S. C. and Glover G. M. C. International journal of chemical reactor engineering. *International Journal of Chemical Reactor Engineering*, 3, 2005.

[159] Zhang H., Zhang K., and Fan S. *CFD* simulation coupled with population balance equations for aerated stirred bioreactors. *Engineering in Life Sciences*, 9(6):421–430, 2009.

[160] Kelly W. J. Using computational fluid dynamics to characterize and improve bioreactor performance. *Biotechnology and Applied Biochemistry*, 49(Pt 4): 225–238, 2008.

[161] Barrue H., Bertrand J., Cristol B., and Xuereb C. Eulerian simulation of dense solid-liquid suspension in multi-stage stirred vessel. *Journal of Chemical Engineering of Japan*, 34(5):585–594, 2001.

[162] Micale G., Montante G., Grisafi F., Brucato A., and Godfrey J. *CFD* simulation of particle distribution in stirred vessels. *Chemical Engineering Research and Design*, 78(3):435–444, 2000.

[163] Decker S. and Sommerfeld M. Calculation of particle suspension in agitated vessels with the euler-lagrange approach. *Institution Of Chemical Engineers Symposium Series*, 140:71–82, 1996.

[164] Sommerfeld M. and Decker S. State of the art and future trends in *CFD* simulation of stirred vessel hydrodynamics. *Chemical Engineering & Technology*, 27(3):215–224, 2004.

[165] Derksen J. J. Numerical simulation of solids suspension in a stirred tank. *AIChE Journal*, 49(11):2700–2714, 2003.

[166] Bezzo F., Macchietto S., and Pantelides C. C. General hybrid multizonal *CFD* approach for bioreactor modeling. *AIChE Journal*, 49(8):2133–2148, 2003.

[167] Elqotbi M., Montastruc L., Vlaev S. D., and Nikov I. Cfd stimulation of gluconic acid production in a stirred gas-liquid fermenter. In *12th European Conference on Mixing Bologna*, pages –, 2006.

[168] Lapin A., Schmid J., and Reuss M. Modeling the dynamics of *E. coli* populations in the three-dimensional turbulent field of a stirred-tank bioreactor – a structured-segregated approach. *Chemical Engineering Science*, 61(14): 4783–4797, 2006.

[169] Avery S. V. Microbial cell individuality and the underlying sources of heterogeneity. *Nature Reviews Microbiology*, 4(8):577–587, 2006.

[170] Howlett N. G. and Avery S. V. Flow cytometric investigation of heterogeneous copper-sensitivity in asynchronously grown *Saccharomyces cerevisiae*. *FEMS Microbiology Letters*, 176(2):379–386, 1999.

[171] Sumner E. R., Avery A. M., Houghton J. E., Robins R. A., and Avery S. V. Cell cycle- and age-dependent activation of sod1p drives the formation of stress resistant cell subpopulations within clonal yeast cultures. *Molecular Microbiology*, 50(3):857–870, 2003.

[172] Kale S. P. and Jazwinski S. M. Differential response to uv stress and dna damage during the yeast replicative life span. *Developmental Genetics*, 18 (2):154–160, 1996.

[173] Jorgensen P. and Tyers M. How cells coordinate growth and division. *Current Biology*, 14(23):R1014–R1027, 2004.

[174] Jorgensen P., Nishikawa J. L., Breitkreutz B.-J., and Tyers M. Systematic identification of pathways that couple cell growth and division in yeast. *Science*, 297(5580):395–400, 2002.

[175] Saldanha A. J., Brauer M. J., and Botstein D. Nutritional homeostasis in batch and steady-state culture of yeast. *Molecular Biology of the Cell*, 15(9):4089–

4104, 2004.

[176] Porro D., Brambilla L., and Alberghina L. Glucose metabolism and cell size in continuous cultures of *Saccharomyces cerevisiae*. *FEMS Microbiology Letters*, 229(2):165–171, 2003.

[177] Rupeš I. Checking cell size in yeast. *Trends in Genetics*, 18(9):479–485, 2002.

[178] Porro D. and Srienc F. Tracking of individual cell cohorts in asynchronous *Saccharomyces cerevisiae* populations. *Biotechnology Progress*, 11(3): 342–347, 1995.

[179] Vanoni M., Vai M., Popolo L., and Alberghina L. Structural heterogeneity in populations of the budding yeast *Saccharomyces cerevisiae*. *Journal of Bacteriology*, 156(3):1282–1291, 1983.

[180] Verduyn C., Postma E., Scheffers W. A., and Van Dijken J. P. Effect of benzoic acid on metabolic fluxes in yeasts: A continuous culture study on the regulation of respiration and alcoholic fermentation. *Yeast*, 8(7):501–517, 1992.

[181] Alberghina L., Rossi R. L., Porro D., and Vanoni M. A modular systems biology analysis of cell cycle entrance into *S*-phase. In Alberghina L. and Westerhoff H., editors, *Systems Biology. Topics in Current Genetics*, volume 13, pages 325 – 347. Springer-Verlag, Berlin/Heidelberg, 2005.

[182] Sonnleitner B. and Käppeli O. Growth of *Saccharomyces cerevisiae* is controlled by its limited respiratory capacity: Formulation and verification of a hypothesis. *Biotechnology and Bioengineering*, 28(6):927–937, 1986.

[183] Morales-Rodriguez R., Meyer A. S., Gernaey K. V., and Sin G. A framework for model-based optimization of bioprocesses under uncertainty: Lignocellulosic ethanol production case. *Computers & Chemical Engineering*, 42(0): 115–129, 2012.

[184] Kumar S. and Ramkrishna D. On the solution of population balance equations by discretization: A fixed pivot technique. *Chemical Engineering Science*, 51(8):1311–1332, 1996.

[185] Nopens I., Beheydt D., and Vanrolleghem P. A. Comparison and pitfalls of different discretised solution methods for population balance models: a simulation study. *Computers & Chemical Engineering*, 29(2):367–377, 2005.

[186] Kumar S. and Ramkrishna D. On the solution of population balance equations by discretization ii. a moving pivot technique. *Chemical Engineering Science*, 51(8):1333–1342, 1996.

[187] Brauer M. J., Saldanha A. J., Dolinski K., and Botstein D. Homeostatic adjustment and metabolic remodeling in glucose-limited yeast cultures. *Molecular Biology of the Cell*, 16(5):2503–2517, 2005.

[188] Cipollina C., Alberghina L., Porro D., and Vai M. Sfp1 is involved in cell size modulation in respiro fermentative growth conditions. *Yeast*, 22(5):385–399, 2005.

[189] Werner-Washburne M., Roy S., and Davidson G. S. Aging and the survival of quiescent and non-quiescent cells in yeast stationary-phase cultures. In Breitenbach M., Jazwinski S. M., and Laun P., editors, *Aging Research in Yeast*, volume 57, pages 123–143. Springer Netherlands, Dordrecht, 2011.

[190] Vanoni M., Rossi R. L., Querin L., Zinzalla V., and Alberghina L. Glucose modulation of cell size in yeast. *Biochemical Society Transactions*, 33(Pt 1): 294–296, 2005.

[191] Youk H. and Oudenaarden A. v. Growth landscape formed by perception and import of glucose in yeast. *Nature*, 462(7275):875–879, 2009.

[192] Talia S. D., Skotheim J. M., Bean J. M., Siggia E. D., and Cross F. R. The effects of molecular noise and size control on variability in the budding yeast cell cycle. *Nature*, 448(7156):947–951, 2007.

[193] Di Talia S., Wang H., Skotheim J. M., et al. Daughter-specific transcription factors regulate cell size control in budding yeast. *PLoS Biology*, 7(10), 2009.

[194] Sidoli F. R., Asprey S. P., and Mantalaris A. A coupled single cell-population-balance model for mammalian cell cultures. *Industrial & Engineering Chemistry Research*, 45(16):5801–5811, 2006.

[195] Mancuso L., Liuzzo M. I., Fadda S., et al. Experimental analysis and modelling of in vitro proliferation of mesenchymal stem cells. *Cell Proliferation*, 42(5):602–616, 2009.

[196] Fadda S., Cincotti A., and Cao G. A novel population balance model to investigate the kinetics of in vitro cell proliferation: Part ii. numerical solution, parameters' determination, and model outcomes. *Biotechnology and Bioengineering*, 109(3):782–796, 2012.

[197] Shu C.-C., Chatterjee A., Dunny G., Hu W.-S., and Ramkrishna D. Bistability versus bimodal distributions in gene regulatory processes from population balance. *PLoS Computational Biology*, 7(8), 2011.

[198] Shu C.-C., Chatterjee A., Hu W.-S., and Ramkrishna D. Modeling of gene regulatory processes by population-mediated signaling: New applications of population balances. *Chemical Engineering Science*, 70:188–199, 2012.

[199] Thattai M. and Van Oudenaarden A. Intrinsic noise in gene regulatory networks. *Proceedings of the National Academy of Sciences*, 98(15):8614–8619, 2001.

[200] Hasty J., McMillen D., and Collins J. J. Engineered gene circuits. *Nature*, 420 (6912):224–230, 2002.

[201] Porro D., Martegani E., Ranzi B. M., and Alberghina L. Oscillations in continuous cultures of budding yeast: A segregated parameter analysis. *Biotechnology and Bioengineering*, 32(4):411–417, 1988.

[202] Richard P. The rhythm of yeast. *FEMS Microbiology Reviews*, 27(4):547–557, 2003.

[203] Vrébel P., Lans R. G.van der , Luyben K. C., Boon L., and Nienow A. W. Mix-

ing in large-scale vessels stirred with multiple radial or radial and axial up-pumping impellers: modelling and measurements. *Chemical Engineering Science*, 55(23):5881–5896, 2000.

[204] Schmalzriedt S., Jenne M., Mauch K., and Reuss M. Integration of physiology and fluid dynamics. In *Process Integration in Biochemical Engineering*, pages 19–68. 2003.

[205] Villadsen J., Nielsen J., and Lidén G. Growth kinetics of cell cultures. In *Bioreaction Engineering Principles*, pages 271–357. Springer US, Boston, MA, 2011.

[206] Schweder T., Krüger E., Xu B., et al. Monitoring of genes that respond to process-related stress in large-scale bioprocesses. *Biotechnology and Bioengineering*, 65(2):151–159, 1999.

[207] Neubauer P. and Junne S. Scale-down simulators for metabolic analysis of large-scale bioprocesses. *Current Opinion in Biotechnology*, 21(1):114–121, 2010.

[208] Delafosse A., Delvigne F., Collignon M.-L., et al. Development of a compartment model based on cfd simulations for description of mixing in bioreactors. *Biotechnologie, Agronomie, Société et Environnement*, 14(S2):517–522, 2010.

[209] Morchain J. and Fonade C. A structured model for the simulation of bioreactors under transient conditions. *AIChE Journal*, 55(11):2973–2984, 2009.

[210] Morchain J., Gabelle J.-C., and Cockx A. Coupling of biokinetic and population balance models to account for biological heterogeneity in bioreactors. *AIChE Journal*, page DOI: 10.1002/aic.13820, 2012.

[211] Câmara M., Rubio J. C., Abrã o A., and Davim J. State of the art on micromilling of materials, a review. *Journal of Materials Science & Technology*, 28(8):673–685, 2012.

[212] Hesselman M. C., Odoni D. I., Ryback B. M., et al. A multi-platform flow device for microbial (co-) cultivation and microscopic analysis. *PLoS ONE*, 7 (5), 2012.

[213] Sternberg C., Christensen B. B., Johansen T., et al. Distribution of bacterial growth activity in flow-chamber biofilms. *Applied and Environmental Microbiology*, 65(9):4108–4117, 1999.

[214] Speizer L., Haugland R., and Kutchai H. Asymmetric transport of a fluorescent glucose analogue by human erythrocytes. *Biochimica et Biophysica Acta - Biomembranes*, 815(1):75–84, 1985.

[215] Oh K.-B. and Matsuoka H. Rapid viability assessment of yeast cells using vital staining with 2-nbdg, a fluorescent derivative of glucose. *International Journal of Food Microbiology*, 76:47–53, 2002.

[216] Lantz R. Quantitative evaluation of numerical diffusion (truncation error). *Society of Petroleum Engineers Journal*, 11(3):–, 1971.

[217] Rothman J. *Advances in Cytometry*. Academic Press, 2010.

[218] Kaufmann B. B., Yang Q., Mettetal J. T., and Oudenaarden A.van . Heritable stochastic switching revealed by single-cell genealogy. *PLoS Biology*, 5(9): –, 2007.

[219] Bakstad D., Adamson A., Spiller D. G., and White M. R. Quantitative measurement of single cell dynamics. *Current Opinion in Biotechnology*, 23(1): 103–109, 2012.

[220] Zenklusen D., Larson D. R., and Singer R. H. Single-rna counting reveals alternative modes of gene expression in yeast. *Nature Structural & Molecular Biology*, 15(12):1263–1271, 2008.

[221] Trcek T., Chao J. A., Larson D. R., et al. Single-mrna counting using fluorescent in situ hybridization in budding yeast. *Nature Protocols*, 7(2):408–419, 2012.

[222] Donachie W. D. The cell cycle of *Escherichia coli*. *Annual review of microbiology*, 47:199–230, 1993.

[223] Nielsen J. Modelling the morphology of filamentous microorganisms. *Trends in Biotechnology*, 14(11):438–443, 1996.

[224] Liu G., Xing , M. , and Han Q. A population-based morphologically structured model for hyphal growth and product formation in streptomycin fermentation. *World Journal of Microbiology and Biotechnology*, 21(8-9):1329–1338, 2005.

[225] Xu M., Chen F., Liu X., Ge W., and Li J. Discrete particle simulation of gas-solid two-phase flows with multi-scale *CPU-GPU* hybrid computation. *Chemical Engineering Journal*, 207-208:746–757, 2012.

[226] Crespo A. C., Dominguez J. M., Barreiro A., Gómez-Gesteira M., and Rogers B. D. Gpus, a new tool of acceleration in *CFD*: Efficiency and reliability on smoothed particle hydrodynamics methods. *PLoS ONE*, 6(6), 2011.

[227] Corrigan A. and Löhner R. Semi-automatic porting of a large-scale *CFD* code to multi-graphics processing unit clusters. *International Journal for Numerical Methods in Fluids*, 69(11):1786–1796, 2012.

[228] Lee V. W., Kim C., Chhugani J., et al. Debunking the 100x *GPU vs. CPU* myth: an evaluation of throughput computing on *CPU* and *GPU*. *SIGARCH Computational Architechture News*, 38(3):451–460, 2010.